



CCIS

The Center for Comparative Immigration Studies
University of California, San Diego

**International Migration, Self-Selection, and the
Distribution of Wages: Evidence from Mexico and the
United States**

By Daniel Chiquiar

University of California, San Diego

Gordon H. Hanson

University of California, San Diego
National Bureau of Economic Research

Working Paper 59
September 2002

**International Migration, Self-Selection, and the Distribution of Wages:
Evidence from Mexico and the United States**

September 2002

Daniel Chiquiar
University of California, San Diego

Gordon H. Hanson
University of California, San Diego and
National Bureau of Economic Research

Abstract. In this paper, we use data from the Mexico and U.S. population censuses to examine who migrates from Mexico to the United States and how the skills and economic performance of these individuals compare to those who remain in Mexico. We test Borjas' negative-selection hypothesis that in poor countries the individuals with the strongest incentive to migrate to rich countries are those with relatively low skill levels. We find that 1) Mexican immigrants, while much less educated than U.S. natives, are on average more educated than residents of Mexico, and 2) were Mexican immigrants in the United States to be paid according to current skill prices in Mexico they would tend to occupy the middle and upper portions of Mexico's wage distribution. These results are inconsistent with the negative-selection hypothesis and suggest, instead, that in terms of observable skills there is intermediate or positive selection of immigrants from Mexico. The results also suggest that migration abroad may raise wage inequality in Mexico.

We thank Kate Antonovics, Julian Betts, Richard Carson, Gary Ramey, James Rauch, Christopher Woodruff and seminar participants at El Colegio de Mexico, the NBER, Stanford University, UC Berkeley, UC San Diego, the University of Toronto, and Yale University for helpful comments.

1. Introduction

During the last three decades, the United States has experienced rising levels of immigration from poor countries.¹ This has made the U.S. labor force larger, younger, and less-educated than it otherwise would have been (Smith and Edmonston, 1997; Borjas, 1999). The shift in the composition of immigrants appears to be due in part to changes in U.S. policy embodied in the 1965 Immigration Act, which relaxed long-standing country-of-origin restrictions on immigrant admissions.

More open immigration policies have allowed a wider cross-section of individuals to move to the United States. In an important strand of literature, Borjas (1987) argues that who migrates to the United States from a particular country will depend on that country's wage distribution. In poor countries, where the returns to education and the dispersion of wages are thought to be relatively high, there will be a "negative selection" of immigrants. Those with the greatest incentive to migrate to the United States will tend to be individuals with below-average skill levels in their home countries. In other rich countries, where returns to education and wage dispersion are thought to be relatively low, there will be "positive selection" of immigrants. Those with the greatest incentive to migrate will tend to be individuals with above-average skill. In support of this selection hypothesis, Borjas (1987, 1995) finds that as sources for U.S. immigration have shifted to poor countries, the economic performance of new immigrants has deteriorated. Relative to earlier cohorts, recent immigrants earn lower wages compared to natives at time of arrival and take longer for their earnings to converge to native levels.² These findings

¹ Europe has also had rising immigration from poor countries. See Boeri, Hanson, and McCormick (2002).

² Identifying changes in the average quality of immigrant cohorts is complicated by a confluence of shocks. Over time, unobserved cohort quality may change, immigrants may through assimilation improve their economic performance relative to natives, and labor-market disturbances may have non-uniform effects on

counter an earlier belief that, irrespective of the country of origin, immigrants tend to be individuals with high potential for earnings growth (Chiswick, 1978).³

In another strand of literature, there have been attempts to link recent changes in the level and composition of immigration to rising U.S. wage inequality. Many studies have found that across U.S. regions immigration inflows are only weakly negatively correlated with wage changes for low-skilled U.S. natives, suggesting that immigration has had little impact on the U.S. wage structure (LaLonde and Topel, 1997; Smith and Edmonston, 1997; Borjas, 1999). Borjas, Freeman, and Katz (1997) argue that cross-area wage regressions of this type identify the wage impact of immigration only under restrictive assumptions. Using factor-content calculations, they find larger effects of immigration on the wages of very low-skilled U.S. natives, but this approach too has been criticized (Deardorff, 2000; Leamer, 2000). While the prior of many observers is that U.S. immigration has raised wage inequality, the evidence remains in dispute.

Largely missing in the discussion of the causes and consequences of U.S. immigration is evidence from source countries. Does evidence on who migrates from poor countries to the United States support the negative-selection hypothesis? Has this migration tended to narrow or widen the distribution of wages in source countries?

In this paper, we use data from the 1990 Mexico population census and data on Mexican immigrants in the 1990 U.S. population census to examine who in Mexico migrates to the United States and how their observable skills and wages compare to those

individuals at different points in the wage distribution. See LaLonde and Topel (1992, 1997), Borjas (1999) and Butcher and DiNardo (2002) for differing approaches in how to deal with this issue.

³ Chiswick (1978) finds that immigrant wages tend to converge to and surpass native wages relatively quickly, but Borjas (1987, 1999) criticizes this use of cross-section data to ask time-series questions. There is other evidence of positive selection of migrants in the form of brain drain from poor countries (Beine et al., 2001; Carrington and Detragiache, 1998) and the internal migration patterns of more and less skilled workers (Borjas, Bronars, and Trejo, 1992; Bound and Holzer, 2000).

who remain at home. Mexico is the largest source country for U.S. immigrants. During the 1990's, the country accounted for 31.3% of U.S. immigration. By 2000 the 7.9 million Mexican immigrants in the United States accounted for 27.7% of the U.S. foreign-born population and were equal to 7.8% of Mexico's total population. Relative to the United States, Mexico has high returns to education and high wage dispersion, making it an ideal candidate in which to look for negative selection of migrants.

There is descriptive literature on Mexican immigrants in the United States that tends to support negative selection. Much of this work is based on surveys at the U.S.-Mexico border or in Mexican villages that send farm laborers to the United States (Massey et al., 1994). Durand et al. (2001), for instance, find that migrants going to the United States have low education levels relative to the average Mexican resident. These survey data, however, are subject to two types of sample selection: the surveyed communities tend to be in rural Mexico, where average schooling is low, and the surveyed migrants are those who have chosen to return to Mexico from abroad.⁴ Little work has used nationally representative samples of Mexican residents and emigrants.⁵

To preview our results, we have three sets of findings. First, Mexican immigrants, while much less educated than U.S. natives, are on average more educated than residents of Mexico. Mexican individuals with 12 to 15 years of schooling are the

⁴ Durand et al. (2001) find that the characteristics of Mexican immigrants in the United States have been stable over time. Using alternative survey data, Cornelius and Marselli (2001) find instead that over time Mexican immigrants have become younger, more likely to be female, more educated, more urban, and more likely to settle in the United States on a permanent basis. The contrast in results from different surveys suggests that sample selection issues are important in the analysis of Mexican immigration.

⁵ One exception is Feliciano (2001), who summarizes average schooling levels for Mexican residents and Mexican immigrants in the United States over the last several decades. She finds that in years prior to 1990, as well, average years of schooling were higher for Mexican immigrants than for Mexican residents.

education group most over-represented in the United States. This is initial evidence against the negative-selection hypothesis for Mexico-to-U.S. migration.

Second, wage differences between the United States and Mexico are decreasing in the age and education of the individual. That U.S.-Mexico wage differentials decrease in education and the probability of Mexico-to-U.S. migration increases with education (over a range of values) suggests that migration costs play a pivotal role in determining who migrates. These costs appear to be lower for the more educated, who may be better able to manage the U.S. immigration process or may be less subject to credit constraints.

Third, under existing patterns of wage setting, Mexican immigrants would tend to fall disproportionately in the middle and upper portions of Mexico's wage distribution. To obtain these results, we extend DiNardo, Fortin, and Lemieux's (1996) framework for constructing counterfactual wage densities. We project the wage distribution that would obtain were Mexican immigrants in the United States paid according to Mexico's current wage structure (adjusting for a possible U.S. undercount of illegal immigrants). Our findings show evidence against negative selection and suggest instead there is intermediate or positive selection of Mexican immigrants in terms of observable skills. By removing workers from the middle of Mexico's wage distribution, Mexico-to-U.S. migration may raise wage dispersion in Mexico. Combined with the Borjas, Freeman, and Katz (1997) result that immigration lowers wages for low-skilled U.S. natives, international migration may raise wage inequality in both sending and receiving countries.

The paper has four additional sections. In section 2, we present a simple model of migrant selection. In section 3, we describe the characteristics and economic performance of residents of Mexico and Mexican immigrants in the United States. In

section 4, we use counterfactual wage densities to test for the positive or negative selection of migrants from Mexico. In section 5, we present concluding remarks.

2. Theory

To motivate the empirical analysis, we present a simple model of migration. Borjas (1987, 1991) uses the Roy (1951) model to show that in poor countries migrants tend to be negatively selected (i.e., they are drawn primarily from the lower half of the skill distribution in their home country). This analysis is based on the assumption that time-equivalent migration costs are constant across individuals. In this case, high returns to schooling and high wage dispersion in poor countries give individuals with below average skills the strongest incentive to migrate abroad. In this section, we show that if migration costs are decreasing in the skill level, migrants from poor countries may be either negatively or positively selected in terms of skill, depending on the magnitude of migration costs and the shape of the skill distribution.

Individuals from Mexico, indexed by 0, choose whether or not to migrate to the United States, indexed by 1. For simplicity, we treat this as a one-time decision, though the extension to a dynamic setting is straightforward (see Sjaastad, 1962; Borjas, 1991). Residents of Mexico face a wage equation given by

$$\ln(w_0) = \mu_0 + \delta_0 s \tag{1}$$

where w_0 is the wage in Mexico, μ_0 is the base wage in Mexico, s is the level of schooling, and δ_0 is the return to schooling in Mexico. We focus on migrant selection in terms of observable skills and, in particular, in terms of schooling. Implicitly, we imagine s to be a random variable (and that there are random components to wage

determination), but for simplicity we leave these features of migration behavior in the background. The appendix adds these details to the analysis. If the population of Mexicans were to migrate to the United States, they would face the wage equation

$$\ln(w_1) = \mu_1 + \delta_1 s \quad (2)$$

where w_1 is the U.S. wage for a migrant from Mexico, μ_1 is the base wage for migrants, and δ_1 is the return to schooling for migrants. Consistent with the fact that Mexico has a relatively scarce supply of skill, we assume that $\delta_0 > \delta_1$, or that the return to schooling is higher in Mexico than in the United States.

Let C be migration costs and let $\pi=C/w_0$ be migration costs in time-equivalent units (the number of labor hours needed to migrate to the United States). Combining (1) and (2), a resident of Mexico will migrate to the United States if

$$\ln(w_1) - \ln(w_0 + C) \approx \ln(w_1) - \ln(w_0) - \pi > 0. \quad (3)$$

Borjas (1987, 1999) assumes that π is constant, implying that all individuals require the same number of labor hours in order to migrate to the United States. This assumption simplifies the analysis, but may not be an accurate reflection of reality. We assume that time-equivalent migration costs decrease with schooling, such that

$$\ln(\pi) = \mu_\pi - \delta_\pi s \quad (4)$$

Why might this be the case? First, individuals migrating legally to the United States must satisfy many bureaucratic requirements, involving extensive paperwork and repeated interactions with U.S. immigration authorities. More-educated individuals may be able to meet these requirements more easily.⁶ Second, a large service industry of

⁶ Over 90% of legal Mexican immigrants in the United States are admitted under family reunification provisions of U.S. immigration law. U.S. admissions criteria give primary status to immediate relatives of U.S. citizens, secondary status to immediate relatives of U.S. legal residents, tertiary status to more distant

lawyers and other specialists exists to help migrants manage the U.S. admissions process. Given the cost of these services is more or less fixed, the time-equivalent cost of migration will be lower for individuals with higher hourly wages. There is also a large service industry oriented towards illegal immigrants (Orrenius, 1999). To enter the United States successfully, an illegal entrant must cross the border, find transport to a safe location in the United States, and obtain counterfeit residency documents (or hire a smuggler to provide these services). These costs are also fixed, implying higher-wage individuals require fewer effective labor hours to migrate to the United States. Third, credit constraints may raise migration costs for low-income individuals, who are also likely to be less educated. Individuals may have to borrow to cover migration costs. If lower-income individuals face higher borrowing costs, due to a higher expected probability of default, they will face higher migration costs.⁷

Combining (3) and (4), Figure 1 shows $\ln(w_0)$ and $\ln(w_1)-\pi$, which defines the cutoff schooling level that determines who migrates to the United States in the case where $\delta_{\pi}=0$ and $\mu_1 - \mu_0 > e^{\mu\pi}$. Here, time-equivalent migration costs are constant and small, which corresponds to the assumptions in Borjas (1987). He focuses on unobservable skills but the analogy to observable skills is straightforward. In Figure 1, there is negative selection of migrants: individuals with schooling less than s^* migrate from Mexico to the United States and individuals with schooling greater than s^* remain in Mexico. Individuals with relatively high levels of schooling are less likely to migrate to the United States due to the fact that the return to schooling is higher in Mexico.

relatives of U.S. citizens, etc. While obtaining legal assistance cannot change an individual's priority status for admissions, it can help an individual clear the queue for legal admission more quickly.

⁷ A related possibility is that more-educated individuals may face less uncertainty with regards to the U.S. wages they would earn, making them more likely to migrate for any given wage differential.

Figure 2 shows an alternative case in which $\delta_\pi > 0$ and $\mu_1 - \mu_0 < e^{\mu\pi}$.⁸ Here, time-equivalent migration costs are decreasing in schooling and are large. Individuals with schooling in the interval (s_L, s_U) migrate from Mexico to the United States and those with schooling outside this interval remain in Mexico. The selection of migrants in terms of observable skills depends on the distribution of schooling in Mexico. There are three possible cases: (a) *negative selection*: if the mass of the schooling distribution is above s_L , migrants will primarily be individuals with low schooling relative to those who remain in Mexico, (b) *positive selection*: if the mass of the schooling distribution is below s_U , migrants will primarily be individuals with relatively high schooling, or (c) *intermediate selection*: if the schooling distribution has mass both below s_L and above s_U , migrants will primarily be individuals with intermediate schooling. In case (c), fixed migration costs preclude those with low schooling from migrating and high returns to schooling in Mexico dissuade those with high schooling from migrating. Individuals with intermediate schooling thus have the strongest incentive to migrate to the United States. In the absence of data on the model's parameters, we cannot *a priori* determine which case is relevant for Mexico. We leave it to the empirical analysis to resolve the issue.

One important caveat is that our analysis addresses observable skills, only. If the correlation between observable skills and unobservable skills is positive and sufficiently strong, we expect that our results would apply to migrant selection in terms of unobservable skills, as well. See the appendix for discussion.

⁸ If $\mu_1 - \mu_0 > e^{\mu\pi}$ (migration costs are small), then even if migration costs are decreasing in schooling there is still an unambiguous prediction for the negative selection of migrants. An additional assumption needed to obtain Figure 2 is that $\partial(\ln w_1 - \pi) / \partial s = \delta_1 + \delta_\pi e^{\mu\pi - \delta_\pi s} > \delta_0$ for small s .

3. Data and Preliminary Evidence

To compare economic outcomes of residents of Mexico to those of immigrants from Mexico, we need information on worker characteristics in both Mexico and the United States. We use the 1% sample of individuals from the 1990 *General Census of Population and Housing* for Mexico and the 5% PUMS (public use microdata sample) from the 1990 *Census of Population and Housing* for the United States.

We are interested in which Mexican individuals choose to migrate to the United States and in how these individuals would perform in Mexico relative to the performance of non-migrants. For the sample of Mexican immigrants in the United States, we focus on recent immigrants, defined as those who migrated within the last 10 years. They reflect the composition of individuals admitted under current U.S. immigration policy.

One important measurement issue is whether we have random samples of individuals in each country. In the United States, a large fraction of immigrants from Mexico are in the country illegally. In 1990, out of the 4.3 million Mexican-born individuals in the United States recorded in that year's census, 1.0 million were estimated to be illegal immigrants (Costanzo *et al.*, 2001). Might we be concerned that illegal immigrants are under-represented in the U.S. population census? Costanzo *et al.* (2001) estimate that of the 19.8 million foreign-born individuals enumerated in the 1990 U.S. population census, 3.8 million were illegal immigrants. They further approximate that the 1990 census undercounted illegal immigrants by 15-20%.⁹ This matters for our analysis if the labor-market characteristics of these individuals differ systematically from

⁹ In 2000, the U.S. Census Bureau estimates that of the 31.1 million foreign-born individuals enumerated in the census for that year, 7 million were illegal immigrants (Robinson, 2001). It approximates that the 2000 census undercounted illegal immigrants by 15%.

those of immigrants in the sample. In section 4, we conduct sensitivity analysis to assess the potential impact of an illegal immigrant undercount on our results.

3.1 Summary Data

We begin the data summary by considering the magnitude of migrant outflows from Mexico to the United States. While the size of the Mexican immigrant population in the United States is well documented, the size of this population relative to the population of Mexico is perhaps less appreciated. Table 1 shows Mexican immigrants in the United States as a percentage of the population of Mexico by age cohort for males and females. The sample is all individuals 16-65 years of age. Among males 25-34 years old, the 1990 stock of Mexican immigrants in the United States was equivalent to 12.0% of that age cohort in Mexico. For females, this figure is 8.2%. The size of the Mexican immigrant population relative to the population of Mexico is smaller for younger and older cohorts. The within-decade outflows of individuals from Mexico are also large. Among males 25-34 years old, the accumulated net outflow of Mexican migrants to the United States during the 1980's was equal to 6.5% of this cohort's 1990 population in Mexico. For females, this figure is 4.3%. Again, the magnitude of migrant outflows relative to the population of Mexico is smaller for younger and older cohorts. Overall, the accumulated outflows of Mexican migrants appear large enough to have substantial labor-market consequences in Mexico.

Tables 2a and 2b show means for age, educational attainment, labor-force participation, and hourly wages for residents of Mexico, all Mexican immigrants in the United States, recent Mexican immigrants in the United States, and, for comparison,

other U.S. immigrants and U.S. natives. We choose education categories that permit direct comparisons between U.S. and Mexico census data.¹⁰ These categories correspond roughly to modes for high grade of schooling completed in Mexico, which occur at grade 6 (primary education), grade 9 (secondary education), and grade 12 (preparatory education). Some immigrants arrive in the United States as children and continue their schooling. They may have moved to the United States as a result of decisions made by other family members and not by their own choice. Their level of schooling may reflect U.S. educational opportunities and so be endogenous to having migrated. To focus on migrants who attended school in Mexico and who made their own migration decisions, we exclude from the migrant sample individuals under age 17 at time of U.S. entry.¹¹

Tables 2a and 2b reproduce the familiar facts that when compared to other U.S. residents Mexican immigrants in the United States are younger, are much less educated, and have much lower hourly wages.¹² In 1990, 72.3% of all Mexican immigrant men and 68.2% of recent Mexican immigrant men had completed 11 or fewer years of school, compared to only 19.0% of U.S. native men. However, Mexican immigrants, and recent immigrants in particular, compare favorably when we examine educational attainment in Mexico. In 1990, 81.0% of male residents of Mexico had 11 or fewer years of schooling.

¹⁰ A natural comparison of educational attainment in Mexico and the United States would be the simple means of schooling years completed. Unfortunately, for individuals with less than 9 years of schooling, the 1990 U.S. Census only reports educational attainment as zero schooling, 1-4 years of schooling, or 5-8 years of schooling. Given the large mass of Mexican immigrants in the United States with less than 9 years of schooling, it is difficult to calculate the simple mean of schooling for this group.

¹¹ The results are unaffected by raising the cutoff age to 19 or 21 years, suggesting the Mexican immigrants most likely to continue their education in the United States are those that arrive as children. Adult immigrants appear likely to continue their education after arriving in the United States. Betts and Lofstrom (1998) find that school enrollment rates for adult immigrants are higher than for adult U.S. natives. However, the same does not appear to hold for immigrants from Mexico (Borjas, 1996; Trejo, 1997).

¹² For Mexico, average hourly wages are calculated as monthly labor income/(4.5*hours worked last week); for the United States, average hourly wages are calculated as annual labor income/(weeks worked last year*usual hours worked per week).

Relative to male residents of Mexico, Mexican immigrant men are less likely to have 11 or fewer years of education, more likely to have 12-15 years of education, and less likely to 16 plus years of education. A similar pattern holds for women.

Table 2 gives preliminary evidence against the negative-selection hypothesis. In terms of observable skills, it is the more educated, not the least educated, who are more likely to migrate. One concern is that recent Mexican immigrants may tend to have high levels of schooling in part because they are relatively young and educational attainment in Mexico has been rising over time. To control for the effects of age on schooling, Table 3 shows average schooling for 25-34 year old Mexican residents and Mexican immigrants in the United States. For this age cohort, which is the largest among Mexican immigrants in the United States, we again see that Mexican immigrants have high education levels relative to Mexican residents.

Returning to Tables 2a and 2b, there appear to be differences in labor-force participation rates between residents of Mexico and Mexican immigrants in the United States. To examine this issue further, Tables 4a and 4b report the fraction of the 1990 population of Mexican residents and of recent Mexican immigrants in the United States with positive labor earnings by age and schooling cells. This definition of labor-force participation reflects the sample of individuals for which we have wage observations. For males 25-55 years of age with more than four years of education, labor-force participation rates in the two countries are quite similar. Participation rates are higher for Mexican immigrant males in the youngest cohort (16-24 years), the oldest cohort (55-65 years), and with little schooling (0-4 years). To minimize selection issues related to the labor-force participation, we focus the analysis on males aged 25-65.

Labor-force participation rates for women, however, differ markedly between migrants and non-migrants. Among women with 11 or fewer years of education, immigrants are much more likely to have positive labor earnings. This could be due to more elastic female labor supply, in which case higher wages in the United States would induce higher rates of labor-force participation. Also, women who are more likely to work at any wage level may be more likely to self-select into migration. In either case, Mexican immigrant women in the United States who work may differ from the subpopulation of these women that would work were they to return to Mexico.¹³

This is a problem for our later empirical analysis. In constructing counterfactual wage densities, we will want to examine a population of Mexican migrants that would be likely to participate in the labor force were they to return to Mexico. The fact that participation rates differ considerably between migrants and non-migrants, and that these differences are larger for specific age and education groups, may affect the pattern of migrant self-selection that we uncover from data on wage-earning individuals and, in particular, may bias the results towards finding negative selection. This potential bias is larger in the case of females, for whom we observe large differences in labor-force participation rates. We return to this issue in section 4.

3.2 Returns to Education and Mexico-U.S. Wage Differentials

To examine the incentive to migrate, we report evidence on differences in wage outcomes for residents of Mexico and Mexican immigrants in the United States. As an initial indication, Figures 3a and 3b show kernel density estimates for the distribution of wages from the samples of Mexican immigrants and Mexican residents. Not

¹³ See Baker and Benjamin (1997) for further discussion of immigrant male and female labor supply.

surprisingly, wages in Mexico are on average much lower and exhibit greater dispersion, as compared to wages for Mexican immigrants in the United States.

To summarize differences in returns to observable skills in the two countries, we estimate OLS wage regressions for five samples of individuals: residents of Mexico, recent Mexican immigrants in the United States, all Mexican immigrants in the United States, other immigrants in the United States, and U.S. natives. Tables 5a and 5b report the results. The regressors are dummy variables for seven categories of educational attainment, age, age squared, a dummy variable for marital status, a dummy variable for residence in a metropolitan area, dummy variables for race, and, for immigrants, dummy variables for year of entry into the United States.

The estimated returns to education for residents of Mexico are substantially higher than for Mexican immigrants or other groups in the United States.¹⁴ Completing 12 years of schooling is associated with an increase in hourly wages of 66.2 log points for men in Mexico but only 12.4 to 18.2 log points for Mexican-born men in the United States. A similar pattern holds for women. The base wage, as shown by the constant term, is substantially higher in the United States than in Mexico. This difference in the base wage appears to account for the large wage differences between the two countries.¹⁵

¹⁴ The coefficient estimates of these wage regressions may be affected by several factors. Estimates of the returns to schooling may be biased by correlation between unobserved ability and schooling. Also, the self-selection of individuals into the labor force or into migration may introduce further biases. This is of particular concern when comparing migrant and non-migrant wage equations. In unreported results, we attempted to address self-selection related to migration. We estimated wage regressions for Mexican-born individuals in Mexico and in the United States, including as an additional regressor the inverse Mills ratio derived from a probit model of the migration decision. However, we have concerns about these estimates. Since we lack an instrument that affects migration but not wages, identification of the coefficient on the inverse Mills ratio is achieved only through the non-linear way in which the other regressors enter into the inverse Mills ratio and so depends entirely on distributional assumptions. Despite these concerns, correcting for self-selection into migration has little effect on the coefficient estimates.

¹⁵ We exercise caution in interpreting the constant term as it captures differences across countries in mean returns to unobserved skills as well as in returns to raw labor. Given education, age, and age squared are included as regressors, the constant term would algebraically correspond to the log wage of an individual

Table 5 also confirms previous results that estimated U.S. returns to education are lower for recent immigrants and Mexican immigrants relative to other immigrants and for immigrants overall relative to U.S. natives (Borjas, 1996 and 1999).

Who migrates to the United States from Mexico is determined in part by expected wage differences between the two countries. Using the estimated returns to observed characteristics, we calculate the predicted difference in log wages at each age and education level.¹⁶ While this is a crude indication of the incentive to migrate, it is useful for generating approximate binational wage differences at a point in time. One issue in constructing these differences is whether to adjust predicted wages for differences in the cost of living between the two countries. If migrants move to the United States permanently, then we would want to make such an adjustment. In fact, many migrants remit a portion of their earnings to family members in Mexico (Woodruff and Zenteno, 2002) or return to Mexico for extended visits (Massey et al., 1996). For these migrants, the relevant price level may be that in Mexico, in which case the nominal difference in U.S.-Mexico wages is the appropriate one to consider. In the interest of producing more conservative estimates of binational wage differences, we adjust Mexican hourly wages to achieve purchasing power parity with the United States, using PPP adjustment factors reported in the Penn World Tables.¹⁷

Figure 4 shows predicted U.S.-Mexico hourly wage differences by age and education level. Wage differences are largest for the young and for the least educated,

with schooling and age equal to zero. Zero schooling is an appropriate reference point, but zero age is not. Using the estimated coefficients, the predicted log wage for a Mexican-born male who is 25 years old and has zero schooling is -0.93 if the individual resides in Mexico and 1.50 if the individual is in the United States and migrated in the last ten years. This is a smaller difference in U.S. and Mexico base wages than what is implied by the constant terms in Table 5a. A similar comment applies to results for females.

¹⁶ The regressions we use to calculate predicted wages in the United States and Mexico include a complete set of interactions between educational attainment, age, and age squared.

but are quite large in absolute value for all groups. For a 30-year old male, the predicted real U.S. wage premium is \$4.07 per hour for an individual with 5-8 years of schooling, \$3.52 for an individual with 12 years of education, and \$2.60 for an individual with 16 or more years of education.¹⁸ Results for women are similar. Wage differences fall with age in most cases due to the fact that in the United States the returns to age tend to be constant across age levels while in Mexico they rise with age.

That U.S.-Mexico real wage differences appear to be large and positive for all age and education levels suggests that migration costs, broadly defined, are what determine who migrates to the United States. If migration costs were a constant fraction of an individual's wage (i.e., constant in time-equivalent units), then we would expect migration rates to be highest for the least educated. But Table 2 shows that this is not the case. Migration rates are highest for Mexican-born individuals with relatively high levels of education (12-15 years of schooling). To reconcile U.S.-Mexico wage differences that fall with the level of education with migration rates that rise with the level of education (at least up to 15 years of schooling), we need migration costs to be a lower fraction of hourly wages for more-educated individuals.

In this section, we have seen evidence that Mexican immigrants in the United States have high education levels relative to residents of Mexico, that the returns to education for Mexican-born individuals are higher in Mexico than in the United States, and that projected U.S.-Mexico wage differentials for a Mexican-born individual are large for all age-education cohorts and decline with age and schooling. We turn next to

¹⁷ In 1990, Mexico's PPP adjusted price level was 43.3% of the price level for the United States.

¹⁸ Stated in percentage terms, these wage differences are, or course, larger (114 log points for an individual with 9 years of schooling, 83 log points for an individual with 12 years of education, and 55 log points for an individual with 16 or more years of education).

see what the observed differences in the composition of Mexican immigrants and residents of Mexico imply about the distribution of wages for these two groups.

4. Migration Abroad and the Distribution of Wages in Mexico

In this section, we compare wage distributions for residents of Mexico and Mexican immigrants in the United States under common skill prices and labor-force participation behavior. This exercise allows us to determine more formally whether there is positive or negative selection of immigrants from Mexico. It also sheds light on how migration abroad might impact the distribution of wages in Mexico. We begin with a simple comparison of wages for Mexican residents and wages for Mexican immigrants in the United States, using predicted values from OLS wage regressions. We then estimate semi-parametrically the distribution of wages that would obtain for Mexican immigrants were their wages set according to skill prices prevailing in Mexico and were they to participate in the labor force in a manner similar to Mexican residents.

Several important qualifications apply to our results. The framework we use does not address how the distribution of unobserved characteristics might influence the distribution of wages. If, holding age and education constant, Mexican immigrants in the United States have low levels of unobserved ability relative to residents of Mexico, then our framework will tend to understate the extent of negative selection. Our framework also fails to address how Mexico-to-U.S. migration affects skill prices in either country. Such general-equilibrium considerations are beyond the scope of this paper. In section 5, we discuss these and other qualifications in more detail.

4.1 Preliminary Evidence

The first exercise involves the following steps: 1) we estimate OLS wage regressions for male and female residents of Mexico, 2) we use the estimated coefficients to predict wages for wage earners in Mexico and we find the quintiles for this sample, 3) we use the estimated coefficients for residents of Mexico to predict the wages that Mexican immigrants in the United States would earn in Mexico, given their observed characteristics, and 4) we use the predicted wages for Mexican immigrants to calculate the fraction of Mexican immigrants that would occupy each of the wage quintiles for Mexican residents. This exercise is a preliminary attempt to determine whether there is negative or positive selection of Mexican immigrants. If Mexican residents and Mexican immigrants have the same distribution of observed characteristics, then roughly 20% of Mexican immigrants will occupy each of the Mexican-resident wage quintiles. If, on the other hand, Mexican immigrants tend to be negatively selected in terms of observable skills, then immigrants will be over-represented in the lower wage quintiles of Mexican residents and under-represented in the upper wage quintiles.

Table 6 shows how Mexican immigrants in the United States are distributed across wage quintiles for residents of Mexico when we predict their wages using OLS wage equations for Mexican residents. We present results for all Mexican immigrants and for recent immigrants. For men, Mexican immigrants are relatively concentrated in the third and fourth highest wage quintiles and recent immigrants are relatively concentrated in the fourth highest quintile. Mexican immigrant men, and recent immigrants in particular, would appear to fall disproportionately in the upper half of the wage distribution for Mexican residents. This finding goes against the negative-selection

hypothesis. In terms of predicted wages, the relative youth of Mexican immigrant men is more than compensated for by their higher education levels.

For women, Mexican immigrants are concentrated in the two lower wage quintiles, which would appear to support the hypothesis that immigrants are negatively selected in terms of observable skills. This is surprising, given the high levels of schooling for Mexican immigrant women relative to Mexican resident women in Table 2b. What may account for low predicted wages for Mexican immigrant women is that women with low schooling are much more likely to work in the United States than in Mexico. The same is not true for men. This results in female Mexican wage earners in the United States being younger and less educated relative to female wage earners in Mexico, even though in the total population immigrant women are more educated.

The results suggest that when projecting wages that Mexican immigrants would earn in Mexico, we should control for differences in labor-force participation between the two countries. If, as in Table 6, we make no such adjustment, then we end up with a sample of Mexican immigrant workers selected on the basis of who is willing to work at U.S. skill prices and a sample of Mexican resident workers selected on the basis of who is willing to work at Mexican skill prices. For the comparison of wage distributions in the two countries to be meaningful, individuals should be selected on the basis of a common labor-force participation criterion. We address this issue in the next section.

4.2 Counterfactual Wage Densities

Wage distributions for Mexican immigrants in the United States and for residents of Mexico may differ because of differences in the distribution of skills between the two

groups or because of differences in the prices of skills in the two labor markets. To examine differences in the distribution of skills between Mexican residents and Mexican immigrants, we compare wage distributions for Mexican migrants and Mexican residents under a common set of skill prices. We compute the counterfactual wage density of Mexican immigrants in the United States, assuming they are paid according to Mexico's wage structure, and compare it to the actual distribution of wages in Mexico. This involves applying techniques developed by DiNardo, Fortin, and Lemieux (1996). In practice, their technique requires taking Mexico's actual wage distribution and re-weighting it according to the observed characteristics of immigrant workers.

The analysis in section 4.1 suggests that the distribution of skills may differ in part between Mexican immigrant workers and Mexican resident workers because different skill prices in the two countries create different patterns of labor-force participation. To compare wage distributions in the two countries, it is necessary to hold constant both skill prices and how skill prices influence labor-force participation. To do so, we extend the techniques in DiNardo, Fortin, and Lemieux. We begin with Mexico's actual wage density, which corresponds to Mexico's wage structure integrated over the distribution of resident-worker characteristics. Then, we re-weight this distribution according to the observed characteristics of *all* Mexican immigrants (and not just those who earn positive wages). This simulates the wage distribution that would prevail given immigrant characteristics and Mexican skill prices and labor-force participation rates.

Let $f^i(w|x)$ be the density of wages w in country i , conditional on a set of observed characteristics x . Also, let D_i be an indicator variable equal to 1 if the individual is in the labor force and equal to zero otherwise. We further define

$h(x|i=Mex, D_i=1)$ as the density of observed characteristics among wage earners in Mexico, and $h(x|i=US, D_i=1)$ as the density of observed characteristics among wage-earning Mexican immigrants in the United States. The observed density of wages for individuals working in Mexico is

$$g(w|i=Mex, D_i=1) = \int f^{Mex}(w|x)h(x|i=Mex, D_i=1)dx \quad (5)$$

Likewise, the observed density of wages for Mexicans working in the U.S. is

$$g(w|i=US, D_i=1) = \int f^{US}(w|x)h(x|i=US, D_i=1)dx \quad (6)$$

Differences in $f^{Mex}(w|x)$ and $f^{US}(w|x)$ capture differences in skill prices in the two countries.¹⁹ Differences in $h(x|i=Mex, D_i=1)$ and $h(x|i=US, D_i=1)$ capture differences in the distribution of observed characteristics for Mexican immigrant workers and for Mexican resident workers. The differences in the $h(\cdot)$ functions are due in part to differences in the characteristics of Mexican immigrants and Mexican residents and in part to differences in who participates in the labor force in the two countries.

Consider the density of wages that would prevail for Mexican immigrant workers in the United States if they were paid according to the prices of skills in Mexico:

$$g_{US}^{Mex}(w) = \int f^{Mex}(w|x)h(x|i=US, D_i=1)dx \quad (7)$$

This corresponds to the distribution of wages for Mexican residents in (5), except that it is integrated over the skill distribution for working Mexican immigrants in the United States. While this distribution is unobserved, we can rewrite it as

¹⁹ When the conditional expectation is linear in the observed characteristics, these terms are closely related to the regression equation for wages on observable characteristics (Butcher and DiNardo, 2002).

$$\begin{aligned}
g_{US}^{Mex}(w) &= \int f^{Mex}(w|x)h(x|i=US, D_i=1) \frac{h(x|i=Mex, D_i=1)}{h(x|i=Mex, D_i=1)} dx \\
&= \int \theta f^{Mex}(w|x)h(x|i=Mex, D_i=1) dx
\end{aligned} \tag{8}$$

where

$$\theta = \frac{h(x|i=US, D_i=1)}{h(x|i=Mex, D_i=1)} \tag{9}$$

The key insight of DiNardo, Fortin, and Lemieux (1996) is that a counterfactual density as in (7) can be estimated by taking an observed density (e.g., for wage earners in Mexico) and re-weighting it (e.g., to reflect the distribution of characteristics of Mexican immigrant workers) as in (8). To compute the weights, use Bayes' Law to write,

$$h(x) = \frac{h(x|i=US, D_i=1) \Pr(i=US, D_i=1)}{\Pr(i=US, D_i=1|x)} \tag{10}$$

and,

$$h(x) = \frac{h(x|i=Mex, D_i=1) \Pr(i=Mex, D_i=1)}{\Pr(i=Mex, D_i=1|x)} \tag{11}$$

Combining (10) and (11) we can obtain an expression for θ that is a function of the ratio of the probability that a Mexican-born individual works in Mexico (conditional on x) to the probability that a Mexican-born individual works in the United States (conditional on x). DiNardo, Fortin, and Lemieux (1996) suggest estimating these probabilities parametrically, using the estimates to calculate θ , and then applying the estimated θ 's to standard kernel density estimators to obtain a counterfactual wage density as in (8).

But the counterfactual density in (8) is not what we desire, precisely. The weight, θ , as seen in (9), adjusts for differences in the distribution of skills between Mexican immigrants and Mexican residents, *conditional* on Mexican immigrants working in the United States and Mexican residents working in Mexico. In order to compare wage

distributions for the two groups, we want to condition on common labor-force participation behavior. This requires modifying the weight we use to construct the counterfactual wage density. First, note that it is possible to write the joint probability of migration and labor-force participation, conditional on x , as the product of the conditional distribution of the participation outcome and the marginal of the migration outcome:

$$\Pr(i = US, D_i = 1 | x) = \Pr(D_i = 1 | i = US, x) \Pr(i = US | x) \quad (12)$$

$$\Pr(i = Mex, D_i = 1 | x) = \Pr(D_i = 1 | i = Mex, x) \Pr(i = Mex | x) \quad (13)$$

Given (9)-(13), we can write:

$$\theta = \frac{\Pr(D_i = 1 | i = US, x) \Pr(i = US | x) \Pr(i = Mex, D_i = 1)}{\Pr(D_i = 1 | i = Mex, x) \Pr(i = Mex | x) \Pr(i = US, D_i = 1)} \quad (14)$$

Next, note that $\Pr(i = Mex, D_i=1)/\Pr(i = US, D_i=1)$ is a constant given by the sample proportions of Mexican resident and Mexican immigrant workers. Since θ is scaled to sum to one, we can without loss of generality set $\Pr(i = Mex, D_i=1)/\Pr(i = US, D_i=1) = 1$. Thus, the weight θ can be written as the product of two ratios:

$$\theta = \theta^P \theta^M \quad (15)$$

where:

$$\theta^P = \left(\frac{\Pr(D_i = 1 | i = US, x)}{\Pr(D_i = 1 | i = Mex, x)} \right) \quad \text{and} \quad \theta^M = \left(\frac{\Pr(i = US | x)}{\Pr(i = Mex | x)} \right) \quad (16)$$

The first ratio – the probability of a Mexican-born individual working, conditional on being in the United States, over the probability of a Mexican-born individual working, conditional on being in Mexico – adjusts Mexico’s wage density in (8) to reflect U.S. labor-force participation rates for each realization of x . The second ratio – the probability of a Mexican-born individual being in the United States over the probability of a

Mexican-born individual being in Mexico – corrects the wage density of Mexican residents to reflect the characteristics of Mexican immigrants.

Clearly, the second ratio, θ^M , is the appropriate weight to construct our desired counterfactual wage density. In (15) and (16), we see that the full weight, θ , adjusts for differences in observed characteristics between Mexican immigrants and Mexican residents *and* for differences in labor-force participation. The density we desire should adjust only for the former set of differences. In (8), we replace θ with θ^M , which yields the wage distribution that would obtain if immigrants were paid according to Mexican skill prices and participated in the labor force as do residents of Mexico.²⁰

To compute θ^M , we estimate $Pr(i=US | x)$ parametrically by running a logit on the probability of a Mexican adult being in the United States using the full sample of Mexican immigrants and Mexican residents (and not just wage earners). Once we estimate this model, we can compute $Pr(i=Mex | x) = 1 - Pr(i=US | x)$ and construct the relevant weight θ_j^M for each observation j in the sample.

Having computed these weights (and normalized them to sum to one), we estimate the counterfactual wage density non-parametrically. A kernel density estimator applied to the sample of wage earners in Mexico, but using the weights described above, would yield an estimate of the desired counterfactual density. We adapt the Rosenblatt-Parzen kernel density estimator to the case in which observations have sample weights,

²⁰ For males, θ and θ^M are close in value (labor-force participation rates among males are similar in the two countries, making θ^p close to one). Using either θ or θ^M to construct the counterfactual wage densities yields similar results. For females, labor-force participation rates differ in the two countries. The counterfactual wage density for females using θ as the weight places greater emphasis on the earnings of women with low levels of schooling and so is more supportive of negative selection (in contrast to the results that we report below using θ^M as the weight). This finding is qualitatively similar to the results in Table 6 and highlights the importance of holding labor-force participation constant.

$$\hat{g}(w) = \sum_{j=1}^n \frac{\theta_j^M}{h} K\left(\frac{w - W_j}{h}\right) \quad (17)$$

where we use $j = 1, \dots, n$ sampled workers in Mexico with wages W_j . In this expression K corresponds to a kernel function, and h is the selected bandwidth.

This methodology detects differences in the wage distributions of Mexican residents and Mexican immigrants attributable to differences in observed characteristics. If observable characteristics understate differences in the true distribution of skills, then the results may underestimate the actual difference in corresponding wage densities.

4.3 Main Results

We apply the methodology to our combined sample of Mexican residents and recent Mexican immigrants in the United States (individuals born in Mexico who are 25-65 years of age). To construct counterfactual wage densities, we estimate a logit model for $Pr(i=US | x)$, using the full sample of individuals. The model links the choice of migrating to the United States to the age and age squared of the individual, dummy variables for schooling and marital status, and interactions of these variables. In the case of females, the model also includes the number of children born to an individual.

Predicted probabilities for the logit model are shown in Figure 5.²¹ The probability of migrating to the United States is non-monotonic in age and education. It first rises with age, peaking at 25 to 30 years, and then declines. This is consistent with the idea that the young, who would enjoy high relative earnings abroad over more years,

²¹ The graphs include predicted probabilities using data for individuals with ages from 16 to 65 years of age. This is done to produce a more complete picture of how migration probabilities change with age. The counterfactuals below are based on logit models using data for individuals 25 to 65 years of age. The resulting predicted probabilities do not differ from those shown in the graphs.

have the most to gain from migration. From zero schooling, the migration probability declines with education up to nine years of schooling and then rises, peaking for individuals with 12 years of schooling, and then declines again. The high migration probability for individuals with zero schooling may reflect the large presence of migrant farm workers among Mexican immigrants in the United States. Farm workers tend to be older and are the occupation group most likely to have zero schooling. The low likelihood of the very educated to migrate may be explained by the fact that differences in returns to education for Mexican immigrants and residents of Mexico are largest for this schooling group. They appear to have the least to gain from migration.

Having estimated logit models for males and females and computed the weights, θ^M , we use the sample of wage-earning Mexican residents to construct kernel density estimates for the hourly wages of individuals residing in Mexico and the counterfactual kernel densities for Mexican immigrants in the United States, assuming they are paid according to Mexican skill prices and they follow the same labor-force participation behavior as Mexican residents. In these estimates, the choice of bandwidth is determined by starting with small values and increasing them gradually until a relatively smooth plot of the estimate is achieved. When comparing densities, the same bandwidth is applied to each. All estimates are based on a Gaussian kernel function.²²

The results for Mexican males are in Figure 6a. Although the counterfactual density for Mexican immigrants is fairly close to the actual density of Mexican residents,

²² We first used Silverman's (1986) optimal bandwidth, which minimizes the mean integrated squared error if the data are Gaussian and a Gaussian Kernel is used. However, the resulting densities appeared to be excessively smoothed and could in fact suffer from bias if the data do not conform to the Gaussian assumption. In order to avoid large bias in our estimates, we instead started with a bandwidth of 0.03 log wage units, and sequentially increased it until the resulting densities looked relatively smooth. This usually implied bandwidths in our final estimates of around 0.07 log wage units.

some clear differences are apparent. Contrary to the negative-selection hypothesis, it is *not* the lowest-wage males who exhibit a disproportionately high tendency to migrate to the United States. Under the negative selection hypothesis, we would expect the counterfactual density to exhibit a larger mass in the lower tail and a smaller mass in the upper tail, when compared with the actual wage distribution of Mexico residents. In contrast, the counterfactual wage density we estimate shows that, for wages up to the peak of Mexico's wage distribution, the counterfactual is very close to the actual Mexican wage density. Moreover, the counterfactual has a larger mass for wage values at the peak of the distribution and above. It is not until high wages are reached that the counterfactual density falls below the actual density of resident males. Thus, the counterfactual tends to support an intermediate selection pattern. The negative-selection hypothesis appears to explain only the comparative behavior of medium and high-wage males. The latter are more likely to stay in Mexico than the former.

It appears that the men with the highest likelihood to migrate abroad would have medium-to-high wage levels in Mexico and that those least likely to migrate would have very high wage levels in Mexico. Were Mexican immigrant males in the United States to return to Mexico and be paid according to Mexico's current wage structure, they would tend to fall disproportionately in the middle and upper half of Mexico's wage distribution. This suggests that migration abroad, by taking mass away from the middle of Mexico's wage distribution, may raise wage inequality in the country.

The results for females appear to be even less supportive of negative selection. In Figure 6b, we see that up to the peak of the wage distribution for females in Mexico, the counterfactual density for immigrant females is consistently below the actual density for

female residents of Mexico. To the right of the peak, the counterfactual density is above the actual density for a large range of wage values and practically never falls below the actual density of Mexico residents. Women who migrate from Mexico to the United States appear to be positively selected from the wage distribution. Comparing this with the results in Table 6 suggests that the apparent over-representation of immigrant females in the lowest wage quintiles in Mexico is a result of the different labor-force participation patterns in the two countries and is not indicative of negative selection of female immigrants. This comparison makes clear the importance of assuming common labor-force participation behavior for our exercise.

Summarizing our results, it appears that, for both men and women, those most likely to migrate to the United States are from the middle and upper portions of Mexico's wage distribution. This is inconsistent with negative selection hypothesis. In terms of observable skills, males exhibit behavior consistent with intermediate selection and females exhibit behavior consistent with positive selection.

4.4 Correcting for the Undercount of Illegal Immigrants

As discussed in section 3, the U.S. census may undercount illegal immigrants by as much as 20%. This may matter for our results, since the undercounted illegal immigrants may have observable characteristics that differ systematically from other Mexican immigrants. One group of illegal immigrants that appear likely to be undercounted by the census are undocumented farm workers, who tend to have very low education levels. If there are enough undercounted migrants and they have low enough education levels, then our results could be biased against finding negative selection of

migrants. We cannot remedy this problem directly, since we lack data on which immigrants are illegal within our sample. However, we do have data from other sources on the characteristics of likely illegal immigrants. We use these data to examine the sensitivity of our results to the potential undercount of illegal immigrants.

We repeat the exercises from the previous section, but now modify the immigrant sample so that the sampling weights for the logit analysis used to construct the counterfactuals correspond to those of a synthetic sample in which there is no undercount of illegal immigrants. We re-weight the original sample so that: i) the new weights reflect an increase in the total Mexican immigrant population in the United States, corresponding to the estimated number of undercounted illegal immigrants, and ii) the re-weighting procedure increases the population size to a greater extent for education groups in which out-of-sample evidence suggests that illegal immigrants are more concentrated.

The sample re-weighting procedure draws on data from Costanzo *et al.* (2001) who estimate that, out of the 4.3 million Mexican-born individuals included in the 1990 U.S. census, around 1 million individuals were illegal immigrants. Thus, if we assume a 20% undercount rate, the total number of Mexican-born illegal immigrants in the United States in 1990 was of around 1.25 million.^{23,24} Under this assumption, the census should have recorded around 250,000 more Mexican-born individuals, implying an increase of around 6% in the Mexican-born population actually recorded. Thus, we re-weight the

²³ Costanzo *et al.* (2001) consider both a low undercount rate of 15% and a high rate of 20%. We chose the high rate to maximize the possible impact undercounted migrants could have on our results.

²⁴ This figure may seem too low. Borjas, Freeman and Lang (1991) estimate a total of 1.8 million illegal Mexican aliens in 1980. The U.S. Immigration and Naturalization Service estimates that in 1996 there were 2.7 Mexican illegal immigrants in the United States (U.S. Immigration and Naturalization Service, 1998). However, it is important to recall that more than 2 million undocumented Mexican immigrants were legalized under the 1986 Immigration Reform and Control Act. Costanzo *et al.* (2001) consider this effect in computing their estimates for 1990. Taking legalization into account, other authors have constructed estimates of the illegal immigrant population that are similar to the figure above (Fix and Passel, 1994).

original sample so that the represented population is increased by this percentage.²⁵ In order to make this increased population have the characteristics of illegal immigrants, we adjust the weights by schooling group. The “new” population is assumed to follow the distribution of immigrants by schooling group presented in Orrenius (1999), which is reproduced in Table 7. This distribution is based on the “Mexican Migration Project”, a household sample conducted in rural communities in the states of Mexico where migration to the United States has been traditionally high (see Durand et al., 1996; Massey et al., 1994). A high fraction of the migrants in this sample reported entering the United States illegally. The schooling levels of this group are much lower than those of Mexican immigrants counted in the U.S. census or of residents of Mexico.²⁶

Figures 7a and 7b show the results of this exercise. For both males and females, we use the methodology described in the previous section, so that the immigrant counterfactual assumes the same labor-force participation rates as those of residents of Mexico. The resulting counterfactual density estimates are very similar to those obtained without the illegal undercount correction and have the same qualitative implications. In particular, they still reflect the intermediate selection pattern for males and the positive selection pattern for females found before. This suggests that census undercounts of illegal immigrants do not matter materially for our results. We would need to assume implausibly high undercount rates in order to change qualitatively our main findings.

²⁵ We perform this procedure separately for males and females. There is an implied increase in the population of males by 4.2% and of females by 8.0%.

²⁶ One drawback of using this distribution is that it corresponds to male heads of household between 15 and 65 years of age. In order to adjust the female migrant sample we need to assume the distribution of schooling is similar to that for males (in census data, mean schooling for males and females is similar). Also, mean schooling for this sample is lower than for other samples with high proportions of illegal migrants (Cornelius and Marselli, 2001; Durand, Massey and Zenteno, 2001). By using this sample, we increase the potential effect of the exercise on our results.

Other results based on samples in which illegal immigrants are over-represented tend to support our findings. Orrenius and Zavodny (2001) use survey data on rural, low-skilled illegal migrants in Mexico and find that among this group migrants tend to come from the middle of the schooling distribution. The least educated tend not to migrate to the United States. They also show that a worsening of economic conditions in Mexico tends to increase the skill levels of immigrants, perhaps due to credit constraints becoming more severe for less skilled individuals during downturns.

4.5 Robustness Checks

We have argued that it is important to hold constant how skill prices affect labor-force participation when comparing wage distributions across workers in different countries. So far, we have used skill prices and labor-force participation in Mexico as the basis for these comparisons. This is a natural choice, in that we are interested in the potential effects of migration abroad on Mexican wages. For purposes of evaluating whether immigrants are negatively or positively selected, it is no less valid to use labor-force participation in the United States as the basis for comparison. As a consistency check on our previous results, we estimate the distribution of wages that would obtain if residents of Mexico were paid according to U.S. skill prices and participated in the labor force as do Mexican immigrants and then compare this to the actual wage distribution for Mexican immigrants. If immigrants from Mexico are negatively selected in terms of observable skills, then we would expect the counterfactual wage density for Mexican residents to be right shifted relative to the actual density for Mexican immigrants.

Following the analysis in section 4.2, the counterfactual wage density we wish to estimate is the wage structure for Mexican immigrants in the United States integrated over characteristics of Mexican residents, or

$$g_{Mex}^{US}(w) = \int \frac{1}{\theta^M} f^{US}(w|x) h(x|i = US, D_i = 1) dx \quad (18)$$

where the appropriate weight is the inverse of θ^M , the weight defined in equation (16). This inverse is the probability of a Mexican-born individual being in Mexico over the probability of a Mexican-born individual being in the United States, where both probabilities are conditional on x . Thus, we can obtain the desired counterfactual wage density by taking the actual wage density of Mexican immigrants in the United States and re-weighting it by the characteristics of all Mexican residents. To estimate this density, we perform the same steps described earlier, which entails estimating a logit for the probability that a Mexican-born individual is in the United States, using the logit results to estimate the weight $1/\theta^M$, and then applying the weights to estimate non-parametrically a kernel density for Mexican immigrant wages.

Figures 8a and 8b report the results. For neither males nor females is there support for the hypothesis that Mexican immigrants are negatively selection in terms of observable skills. For males, the results mirror those in Figure 6a. The actual wage density for immigrants is similar to the counterfactual density for residents. The differences that do appear are in the upper half of the wage densities. Compared to the actual density for immigrants, the counterfactual density for resident males shows less mass in the region just to the right of the peak and more mass in the upper tail. Resident males appear to be slightly under-represented in the upper middle portion of the immigrant wage distribution and slightly over-represented in the upper tail. This is

consistent with modest intermediate selection of male immigrants in terms of observable skills. Differences in the actual and counterfactual wage densities in Figure 8a are smaller than those in Figure 6a in part because U.S. returns to education are lower than those in Mexico. This compresses any differences in wage distributions between immigrants and residents that are due to differences in educational attainment. For women, Figure 8b mirrors Figure 6b, with differences in the actual and counterfactual wage densities also being less pronounced. Compared to the actual wage density for immigrants, the counterfactual density for resident females shows more mass to the left of the peak of the distribution and less mass to the right of the peak. This is consistent with modest positive selection of immigrant females in terms of observable skills.

5. Discussion

In this paper, we combine two representative samples of Mexican-born individuals, one of residents of Mexico and another of immigrants in the United States. We use these data to test Borjas' (1987) negative-selection hypothesis that in poor countries the less-skilled are those most likely to migrate to rich countries.

Our results are inconsistent with the negative selection of Mexican immigrants in the United States. For Mexican-born males, there is evidence of intermediate selection and for Mexican-born females there is evidence of positive selection. Those most likely to migrate are young adults with moderately high levels of schooling. Were these immigrants to return to Mexico, they would tend to fall disproportionately in the middle or upper portions of Mexico's wage distribution. This suggests that migration to the

United States may reduce the relative supply of middle and upper-middle wage earners in Mexico, contributing to a widening of earnings inequality in the country.

To arrive at these results, we construct counterfactual wage densities for Mexican immigrants in the United States, in which we use their observed characteristics to project the wages they would earn were they to return to Mexico and be paid according to current patterns of wage determination in the country. These counterfactuals assume that the labor-force participation of immigrants mimics that of residents of Mexico. The results appear robust to a possible undercount of illegal immigrants in the U.S. census.

There are several important caveats to our results. First, we cannot say how the return of Mexican immigrants in the United States would affect Mexico's wage structure. This is a complicated general-equilibrium question that we do not address. We construct counterfactual wage densities to test Borjas' selection hypothesis for Mexico and to gauge from where in Mexico's wage distribution migrants bound for the United States are most likely to be drawn. Second, our findings against negative selection and in favor of intermediate or positive selection are in terms of observable characteristics only. We cannot say whether the unobserved skills of Mexican immigrants in the United States are high or low relative to those who remain in Mexico. Still, it is natural to expect that education is positively correlated with individual ability. In this case, accounting for unobserved skill would be unlikely to change our results drastically.

That immigrants from Mexico appear to be relatively high-wage individuals is surprising. Projected U.S.-Mexico wage differentials are largest for the least educated, suggesting that it is low-wage individuals from Mexico who would have most to gain from migrating to the United States. We appeal to heterogeneity in migration costs as a

way to reconcile migration probabilities that rise with education (at least over a range of schooling levels) with binational wage differentials that fall with education.

The more educated may be better able to negotiate the process of migrating abroad, they may have better access to migration networks that would help them get established in the United States, or they may be less subject to credit constraints in financing migration. It may also be the case that higher discount rates or greater risk aversion among low-skilled individuals make migration abroad relatively less attractive for this group. Finally, a “Washington apples” effect (Alchian and Allen, 1964) may also contribute to positive selection of migrants. Given long queues for U.S. legal admission, two-thirds of Mexican-born individuals entering the country appear to do so illegally on their first attempt. Enforcement against illegal immigrants may act as a head tax, disproportionately penalizing the less skilled. If more-skilled and less-skilled illegal immigrants compete for jobs in the United States, it will be efficient, given fixed migration costs, to fill jobs with more skilled individuals first. While it is plausible that more-educated individuals face lower migration costs, we are unaware of data that would allow us to test this idea. We leave this for future research.

Appendix

This appendix reinterprets the model in the text in terms of Roy (1951). Residents of Mexico face a wage distribution in country $i=0$ (Mexico), 1 (U.S.) given by

$$\ln(w_i) = \mu_i + \delta_i s + \varepsilon_i \quad (\text{A1})$$

where w_i is the wage in country i , μ_i is the zero-schooling mean wage in i , s is the level of schooling, δ_i is the return to schooling in i , and ε_i captures deviations from mean earnings (e.g., returns to unmeasured ability) in i and has distribution $N(0, \sigma_i^2)$. ε_0 and ε_1 have correlation coefficient $\rho_{01} > 0$ (employers in both countries value similar characteristics). Schooling is a random variable with distribution

$$s = \mu_s + \varepsilon_s \quad (\text{A2})$$

where μ_s is mean schooling and $\varepsilon_s \sim N(0, \sigma_s^2)$.

Combining (A1)-(A2), a resident of Mexico will migrate to the United States if

$$\ln\left(\frac{w_1}{w_0 + C}\right) \approx (\mu_1 - \mu_0 - \pi) + \mu_s(\delta_1 - \delta_0) + (\varepsilon_1 - \varepsilon_0) + \varepsilon_s(\delta_1 - \delta_0) > 0 \quad (\text{A3})$$

where C is migration costs and $\pi = C/w_0$ is time-equivalent migration costs. Redefining π slightly relative to the model presented in the text, we assume that time-equivalent migration costs are a random variable with distribution

$$\pi = \mu_\pi + \varepsilon_\pi \quad (\text{A4})$$

where μ_π denotes mean migration costs and $\varepsilon_\pi \sim N(0, \sigma_\pi^2)$, as in Borjas (1991). The correlation coefficient for ε_π and ε_i is $\rho_{i\pi}$ $i=0,1,s$. By (A4), the probability that an individual migrates to the United States is given by

$$\Pr(v > -[\mu_1 - \mu_0 - \mu_\pi + \mu_s(\delta_1 - \delta_0)]) = 1 - \Phi(z) \quad (\text{A5})$$

where $\Phi(z)$ is the standard normal distribution function, $v = (\varepsilon_1 - \varepsilon_0 - \varepsilon_\pi) + \varepsilon_s(\delta_1 - \delta_0)$ and $z = -[\mu_1 - \mu_0 - \mu_\pi + \mu_s(\delta_1 - \delta_0)]/\sigma_v$. The probability above gives the Mexico-to-U.S. migration rate.

Will migrants from Mexico to the United States tend to be individuals with relatively high or low education levels? Given (A1)-(A5) and letting $v' = v/\sigma_v$, the expected level of schooling for a Mexican migrant in the United States is

$$E(s | v' > z) = \mu_s + E(\varepsilon_s | v' > z) = \mu_s + \left(\frac{\sigma_s^2}{\sigma_v} (\delta_1 - \delta_0) - \frac{\sigma_\pi \sigma_s}{\sigma_v} \rho_{s\pi} \right) \lambda(z) \quad (\text{A6})$$

where $\lambda(z)$ denotes the inverse Mills ratio $\phi(z)/[1 - \Phi(z)]$, with $\phi(z)$ being the standard normal density function. Migrants from Mexico will have above (below) average schooling relative to residents of Mexico if the term in the brackets on the right of (A6) is positive (negative). Given $(\delta_1 - \delta_0) < 0$, migrants will have below-average schooling if $\rho_{s\pi}$, the correlation between schooling and time-equivalent migration costs, is not too negative, and above-average schooling if $\rho_{s\pi}$ is negative and large in absolute value relative to $|\delta_1 - \delta_0|$. While we cannot sign the term in brackets in (A6), evidence presented in the text suggests that there is a plausible case to be made for $\rho_{s\pi}$ being negative and so for migrants having above-average education levels. Positive selection of migrants is more likely the stronger is the negative correlation between observable skill and migration costs and the smaller are the excess returns to education in the source country.

By conditioning wages on observable correlates of skill, the analysis above can be extended to account for self-selection in terms of unobservable skills. The implications are similar to those of the model described above. In particular, if earnings dispersion is wider in Mexico than in the United States, as it is apparently the case, this would tend to cause migrants to be negatively selected in terms of unobservable skills. However, positive selection emerges if individuals with higher unobserved skills have lower time-equivalent migration costs and this correlation is sufficiently strong.

References

- Alchian, Armen, and William Allen. 1964. *University Economics*. Belmont, CA: Wadsworth.
- Baker, Michael, and Dwayne Benjamin. 1997. "The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations." *American Economic Review* 87: 705-729.
- Boeri, Tito, Barry McCormick, and Gordon H. Hanson. 2002. *Immigration Policy and the Welfare State*. Oxford: Oxford University Press, forthcoming.
- Beine, Michel, Frederic Docquier, and Hillel Rapoport. 2001. "Brain Drain and Economic Growth." *Journal of Development Economics* 64(1): 275-289.
- Betts, Julian and Magnus Lofstrom. 1998. "The Educational Attainment of Immigrants: Trends and Implications." NBER Working Paper No. 6757.
- Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77(4): 531-553.
- Borjas, George, J. 1991. "Immigration and Self-Selection." In John Abowd and Richard Freeman, eds., *Immigration, Trade, and the Labor Market*. Chicago: University of Chicago Press, pp. 29-76.
- Borjas, George J. 1995. "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" *Journal of Labor Economics* 13(2): 201-245.
- Borjas, George J. 1996. "The Earnings of Mexican Immigrants in the United States." *Journal of Development Economics* 51: 69-98.
- Borjas, George J. 1999. "The Economic Analysis of Immigration." In Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam: North-Holland, pp. 1697-1760.
- Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-Selection and Internal Migration in the United States." *Journal of Urban Economics* 32(2): 159-185.

Borjas, George J., Richard B. Freeman, and Lawrence F. Katz. 1997. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity* 1: 1-90.

Bound, John, and Harry J. Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics* 18(1): 20-54.

Butcher, Kristin F. and John DiNardo. 2002. "The Immigrant and Native-Born Wage Distributions: Evidence from United States Censuses." *Industrial and Labor Relations Review*, forthcoming.

Carrington, William J. and Detragiache, Enrica. 1998. "How Big is the Brain Drain?" IMF Working Paper WP/98/102.

Chiswick, Barry. 1978. "The Effect of Americanization on the Earnings of Foreign-born Men." *Journal of Political Economy* 86(5): 897-921.

Costanzo, Joe, Cynthia Davis, Caribert Irazi, Daniel Goodkind and Roberto Ramirez. 2001. "Evaluating Components of International Migration: The Residual Foreign Born Population." U.S. Bureau of the Census, Division Working Paper No. 61.

Cornelius, Wayne A. and Enrico A. Marselli. 2001. "The Changing Profile of Mexican Migrants to the United States: New Evidence from California and Mexico." *Latin American Research Review*, forthcoming.

Deardorff, Alan V. 2000. "Factor Prices and the Factor Content of Trade Revisited: What's the Use?" *Journal of International Economics* 50(1): 73-90.

DiNardo, John, M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-1044.

Durand, Jorge, William Kandel, Emilio Parrado and Douglas S. Massey. 1996. "International Migration and Development in Mexican Communities." *Demography* 33(2): 249-264.

Durand, Jorge, Douglas S. Massey, and Rene M. Zenteno. 2001. "Mexican Immigration in the United States." *Latin American Research Review* 36(1): 107-127.

Feliciano, Zadia. 2001. "The Skill and Economic Performance of Mexican Immigrants from 1910 to 1990." *Explorations in Economic History* 38: 386-409.

Fix, Michael and Jeffrey Passel. 1994. "Immigration and Immigrants: Setting the Record Straight". Urban Institute.

Hanson, Gordon H., and Antonio Spilimbergo. 1999. "Illegal Immigration, Border Enforcement and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border," *American Economic Review*, 89: 1337-57.

LaLonde, Robert and Robert Topel. 1992. "The Assimilation of Immigrants in the U.S. Labor Market." In George J. Borjas and Richard Freeman, eds, *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, (Chicago, IL: University of Chicago Press): 67-92.

LaLonde, Robert and Robert Topel. 1997. "Economic Impact of International Migration and Migrants." In Mark R. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, (Amsterdam: Elsevier Science): 799-850.

Leamer, Edward E. 2000. "What's the Use of Factor Contents?" *Journal of International Economics* 50(1): 17-50.

Massey, Douglas S., L. Goldring, and Jorge Durand. 1994. "Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities." *American Journal of Sociology* 99(6): 1492-1533.

Orrenius, Pia M. 1999. "Family Networks, Coyote Prices, and the Rural Economy in Migration from Western Mexico: 1965-1994." Mimeo, Federal Reserve Bank of Dallas.

Orrenius, Pia M. and Madeline Zavodny. 2001. "Self-Selection among Undocumented Immigrants from Mexico." Mimeo, Federal Reserve Bank of Dallas.

Robinson, J. Gregory. 2001. "ESCAP II: Demographic Analysis Results." U.S. Census Bureau, <http://www.census.gov/dmd/www/ReportRec2.htm>.

Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3: 135-146.

Silverman, B. 1986. *Density Estimation for Statistics and Data Analysis*. London, Chapman and Hall.

Sjaastad, Larry. A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy*, 70(1), pp. 80-93

Smith, James and Barry Edmonston. 1997. *The New Americans: Economic, Demographic and Fiscal Effects of Immigration*. National Academy Press: Washington, D.C.

Trejo, Stephen J. 1997. "Why do Mexican Americans Earn Low Wages?" *Journal of Political Economy* 105(6): 1235-1268..

Woodruff, Christopher and Rene M. Zenteno. 2001. "Remittances and Microenterprises in Mexico." Mimeo, UC San Diego.

Table 1: Share of U.S. Immigrants from Mexico in Population of Mexico, 1990
(percent)

Age	All Mexican Immigrants		Arrivals 1981-1990		Arrivals 1986-1990	
	Males	Females	Males	Females	Males	Females
16-24	7.53	4.88	5.79	3.32	3.92	2.26
25-34	12.02	8.18	6.46	4.27	2.92	2.10
35-44	10.17	8.16	2.55	2.02	1.21	0.90
45-54	7.57	6.99	1.56	1.40	0.79	0.70
55-65	5.96	6.11	0.83	1.06	0.44	0.56

Notes: This table shows Mexican immigrants in the United States (either all immigrants or immigrants that arrived during the period 1981-1990) as a percentage of the 1990 population of Mexico by age and sex categories.

Table 2a: Summary Statistics for Males, 1990

Variable		Mexico Residents	U.S. Residents			U.S. Natives
			Recent Mex. Immigrants	All Mex. Immigrants	Other Immigrants	
Age		32.7	29.3	35.0	40.5	37.2
Highest Grade of Schooling Completed (%)	0	11.6	11.9	12.8	3.8	0.6
	1 to 4	19.1	10.7	14.6	2.7	0.5
	5 to 8	28.6	29.0	30.9	9.2	3.9
	9	13.9	9.4	7.9	2.9	3.3
	10 to 11	7.7	7.2	6.1	4.7	10.7
	<12	81.1	68.2	72.3	23.2	19.0
	12	6.5	19.9	17.0	24.1	32.4
	13 to 15	4.7	8.5	7.8	21.5	27.7
	16+	7.7	3.5	2.9	31.2	20.9
Wage Earners (%)		66.1	83.0	83.7	78.6	80.9
Hourly wage (standard dev.)		1.45 (2.5)	6.95 (5.73)	8.12 (6.6)	13.92 (12.0)	14.18 (12.3)
N		211,133	36,119	66,660	189,237	708,727

Notes: The sample is all individuals 16-65 years old. Residents of Mexico are a 1% random sample of the *XII Censo General de Poblacion y Vivienda, 1990*; Mexican and other immigrants are drawn from the 1990 5% U.S. PUMS; and U.S. natives are drawn from the 1990 1% U.S. PUMS. Immigrants in the United States are restricted to be individuals 17 years or older at time of entry into the country. Recent immigrants from Mexico are individuals who entered the United States during the period 1981-1990. Schooling variables show the percentage of individuals whose high grade completed is that indicated and wage earners shows the percentage of individuals with positive labor earnings in 1990. Wage levels are in 1990 U.S. dollars. Average wages in Mexico are for those individuals with average hourly earnings greater than \$0.1 and less than \$20 and in the United States are for those individuals with average hourly earnings greater than \$1 and less than \$100. The sample size is all individuals (not just those with valid wages).

Table 2b: Summary Statistics for Females, 1990

		Mexico Residents	U.S. Residents			U.S. Natives
			Recent Mex. Immigrants	All Mex. Immigrants	Other Immigrants	
Age		32.7	31.0	37.1	41.9	37.7
Highest Grade of Schooling Completed (%)	0	15.6	12.3	12.8	4.4	0.5
	1 to 4	19.6	12.0	14.6	3.3	0.3
	5 to 8	28.1	30.0	32.6	10.2	3.1
	9	12.3	8.5	7.3	3.1	2.9
	10 to 11	7.4	6.1	5.3	4.9	10.5
	<12	83.3	68.9	72.6	25.8	17.2
	12	7.6	19.5	17.3	30.3	35.6
	13 to 15	4.5	8.2	7.5	22.2	29.6
	16+	4.6	3.4	2.6	21.7	17.5
Wage Earners (%)		20.7	44.7	48.2	60.5	70.3
Hourly wage (standard dev.)		1.33 (2.1)	6.05 (5.38)	6.62 (5.8)	10.15 (8.4)	9.47 (7.4)
N		228,964	26,643	51,918	213,020	737,267

Notes: See notes to Table 2a.

Table 3: Educational Attainment for Mexico-Born 25-34 Year Olds, 1990

		Mexico Residents	Recent Mex. Immigrants	Mexico Residents	Recent Mex. Immigrants
		Males	Males	Females	Females
Age		29.1	28.8	29.2	28.5
Highest Grade of Schooling Completed (%)	0	7.8	10.7	10.9	10.2
	1 to 4	16.0	9.7	18.5	10.2
	5 to 8	29.2	29.3	30.2	31.7
	9	13.4	8.6	11.4	8.7
	10 to 11	6.0	6.6	5.5	5.6
	<12	72.5	64.9	76.8	66.3
	12	8.4	20.7	9.1	19.6
	13 to 15	6.4	10.1	6.2	9.6
	16+	12.7	4.2	7.9	4.5
N		61,018	12,317	55,436	16,713

Notes: This table shows the fraction of 25-34 year-old individuals who have completed given levels of schooling for residents of Mexico and recent Mexican immigrants in the United States (individuals who arrived during the period 1981-1990). See notes to Table 2a for more details on the sample.

Table 4a: Labor-Force Participation Rates, Males in 1990

Years of Schooling	Residents of Mexico					Recent U.S. Immigrants from Mexico				
	Age Category					Age Category				
	16-24	25-34	35-44	45-54	55-65	16-24	25-34	35-44	45-54	55-65
0	63.0	72.5	73.9	71.8	62.5	76.2	85.4	85.9	84.4	71.3
1 to 4	70.9	80.4	80.8	76.6	64.7	80.5	89.0	89.3	84.3	73.8
5 to 8	66.3	82.7	84.2	79.8	61.8	83.1	89.9	88.8	86.8	72.5
9	62.3	85.7	85.9	80.5	61.6	80.8	89.1	89.7	83.4	69.8
10 to 11	31.1	86.6	88.8	79.4	63.5	79.2	89.5	87.6	83.9	75.8
12	52.5	85.2	88.8	83.9	62.0	82.7	90.2	90.3	88.4	72.1
13 to 15	32.5	82.4	90.3	84.5	64.0	83.3	91.6	91.4	84.6	81.0
16+	51.4	84.4	91.6	87.1	69.2	77.6	89.7	93.6	88.3	89.7

Notes: This table shows the percentage of the population that reported positive labor earnings in 1990 by age category and by highest year of schooling completed for residents of Mexico and for recent Mexican immigrants in the United States (individuals who arrived in the country during the period 1981-1990). Immigrants are individuals aged 17 years or older at time of entry in the United States.

Table 4b: Labor-Force Participation Rates, Females in 1990

Years of Schooling	Residents of Mexico					Recent U.S. Immigrants from Mexico				
	Age Category					Age Category				
	16-24	25-34	35-44	45-54	55-65	16-24	25-34	35-44	45-54	55-65
0	9.7	9.1	9.8	9.1	7.3	34.8	41.6	48.8	39.5	15.4
1 to 4	13.3	10.6	11.5	9.9	7.4	33.2	46.5	54.3	45.6	25.5
5 to 8	18.7	16.3	19.3	16.8	10.5	38.5	48.7	52.5	48.7	29.2
9	27.5	27.4	30.2	26.9	12.1	40.5	50.0	55.1	54.5	37.6
10 to 11	17.9	35.4	40.1	34.8	16.9	40.4	48.5	50.3	63.0	38.3
12	37.2	47.1	48.1	42.0	22.5	46.9	52.1	57.9	50.2	27.4
13 to 15	27.3	55.2	57.6	51.0	27.3	49.8	55.4	58.0	60.7	41.6
16+	47.8	58.9	61.5	54.5	32.9	41.7	55.0	65.1	54.9	34.4

Notes: See notes to Table 4a.

Table 5a: OLS Wage Regressions for Men, 1990

Variable		Mexican Residents	Mexican Immigrants Recent	Mexican Immigrants All	Other Immigrants	U.S. Natives
Constant		-1.748 (0.039)	1.228 (0.080)	1.033 (0.057)	0.959 (0.034)	0.198 (0.021)
High Grade Completed	1 to 4	0.123 (0.009)				
	5 to 8	0.267 (0.009)	0.041 (0.012)	0.068 (0.008)	0.028 (0.010)	0.051 (0.016)
	9	0.420 (0.011)	0.101 (0.018)	0.142 (0.012)	0.043 (0.014)	0.143 (0.017)
	10 to 11	0.547 (0.014)	0.052 (0.020)	0.124 (0.014)	0.106 (0.012)	0.184 (0.016)
	12	0.662 (0.013)	0.124 (0.014)	0.182 (0.010)	0.183 (0.009)	0.328 (0.015)
	13 to 15	0.857 (0.014)	0.198 (0.018)	0.266 (0.012)	0.307 (0.009)	0.438 (0.015)
	16+	1.133 (0.011)	0.391 (0.026)	0.471 (0.022)	0.637 (0.009)	0.716 (0.016)
Age		0.044 (0.002)	0.015 (0.004)	0.021 (0.003)	0.044 (0.002)	0.068 (0.001)
Age ² /100		-0.045 (0.002)	-0.017 (0.005)	-0.027 (0.003)	-0.048 (0.002)	-0.066 (0.001)
Imm. Cohort	1981-84		0.122 (0.009)	0.123 (0.009)	0.094 (0.005)	
	1975-80			0.256 (0.009)	0.212 (0.006)	
	1971-74			0.349 (0.011)	0.321 (0.007)	
	1965-70			0.430 (0.015)	0.395 (0.008)	
	pre 1965			0.503 (0.017)	0.473 (0.008)	
R Sqd.		0.235	0.045	0.112	0.244	0.213
N		98,139	19,940	42,514	134,403	450,756

Table 5b: OLS Wage Regressions for Women, 1990

Variable		Mexican Residents	Mexican Immigrants Recent	Mexican Immigrants All	Other Immigrants	U.S. Natives
Constant		-1.684 (0.069)	1.350 (0.130)	1.213 (0.083)	1.328 (0.036)	0.835 (0.025)
High Grade	1 to 4	0.092 (0.019)				
	5 to 8	0.217 (0.017)	0.002 (0.017)	0.030 (0.011)	0.011 (0.010)	-0.019 (0.021)
	9	0.459 (0.020)	0.046 (0.026)	0.080 (0.017)	0.026 (0.013)	0.002 (0.022)
	10 to 11	0.615 (0.023)	0.039 (0.030)	0.084 (0.022)	0.057 (0.012)	0.041~ (0.020)
	12	0.758 (0.019)	0.111 (0.020)	0.154 (0.013)	0.146 (0.008)	0.214 (0.020)
	13 to 15	0.906 (0.020)	0.176 (0.027)	0.253 (0.017)	0.338 (0.009)	0.396 (0.020)
	16+	1.078 (0.019)	0.355 (0.039)	0.446 (0.031)	0.625 (0.009)	0.736 (0.020)
	Age	0.043 (0.003)	0.002 (0.007)	0.008 (0.004)	0.013 (0.002)	0.034 (0.001)
Age ² /100	-0.044 (0.004)	0.000 (0.009)	-0.009 (0.005)	-0.016 (0.002)	-0.033 (0.001)	
Imm. Cohort	1981-84		0.060 (0.013)	0.064 (0.013)	0.113 (0.006)	
	1975-80			0.146 (0.014)	0.221 (0.007)	
	1971-74			0.216 (0.016)	0.282 (0.007)	
	1965-70			0.245 (0.020)	0.319 (0.008)	
	pre 1965			0.272 (0.023)	0.335 (0.008)	
R Sqd.	0.230	0.032	0.061	0.182	0.176	
N	29,193	8,689	20,560	117,187	404,792	

Notes to Table 5:

Reported coefficients are for OLS regressions using the log average hourly wage as the dependent variable. Heteroskedasticity-consistent standard errors are in parentheses. The samples are individuals 25-65 years of age who are residents of Mexico (column 1), Mexican immigrants in the United States who arrived during the period 1981-1990 (column 2), all Mexican immigrants in the United States (column 3), other (non-Mexico-born) immigrants in the United States (column 4), or native-born U.S. residents (column 5). Immigrant samples exclude individuals aged 16 years or less at time of entry in the United States. The Mexican-resident sample excludes individuals with hourly wages less than \$0.10 or greater than \$20 and U.S.-resident samples exclude individuals with hourly wages less than \$1 or greater than \$100 (monetary units are 1990 dollars). Additional regressors (not shown) are dummy variables for marital status and residence in a metropolitan area (all columns), dummy variables for race (columns 4 and 5), and dummy variables for year of entry in the United States (columns 2-4).

**Table 6:
Predicted Wages for Mexican Immigrants and Wage Quintiles in Mexico**

Wage Quintile for Residents of Mexico	% of Mex. Immigrants with Predicted Wages in Quintile			
	Men		Women	
	All Immigrants	Recent Immigrants	All Immigrants	Recent Immigrants
1	18.6	19.1	19.3	23.1
2	17.2	17.8	29.6	28.5
3	21.8	20.2	20.3	19.0
4	23.2	26.4	17.5	19.2
5	19.2	16.5	13.3	10.2

This table shows how Mexican immigrants in the United States would be distributed across wage quintiles for residents of Mexico, where the quintiles go from lowest to highest earnings. Quintiles are constructed using an OLS wage equation estimated on residents of Mexico to predict within sample wages. The same OLS wage equation is then used to predict wages for Mexican immigrants in the United States. See the text for details on the estimation.

Table 7: Distribution of the Mexican Migration Project Sample by Schooling

	Migrants	Non-migrants
Schooling (%)		
0-1 years	27.1	23.3
2-4 years	32.6	22.1
5-6 years	23.1	22.5
7-9 years	8.9	11.7
10-12 years	4.3	8.1
13 or more years	4.1	12.3
% of trips to U.S. involving illegal entry	60.8	---

Notes: This table is reproduced from Orrenius (1999). The sample consists of head of household males between 15 and 65 years of age, surveyed between 1987-1995 in 35 communities in western Mexico.

Figure 1

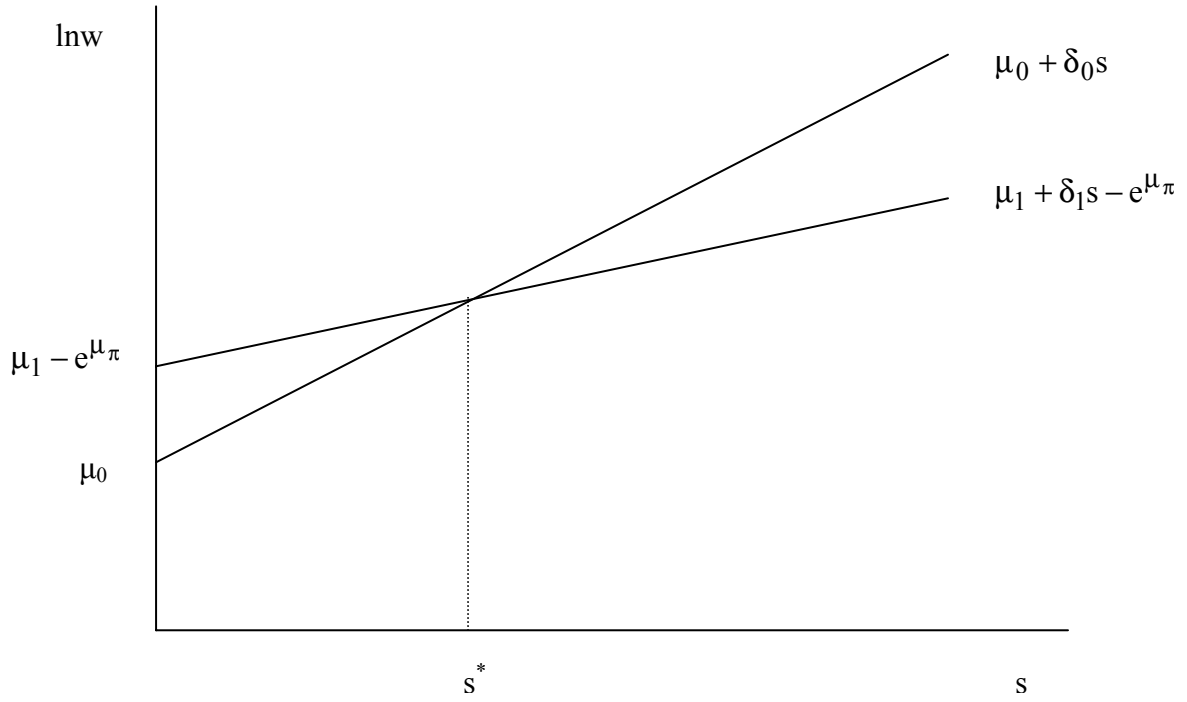


Figure 2

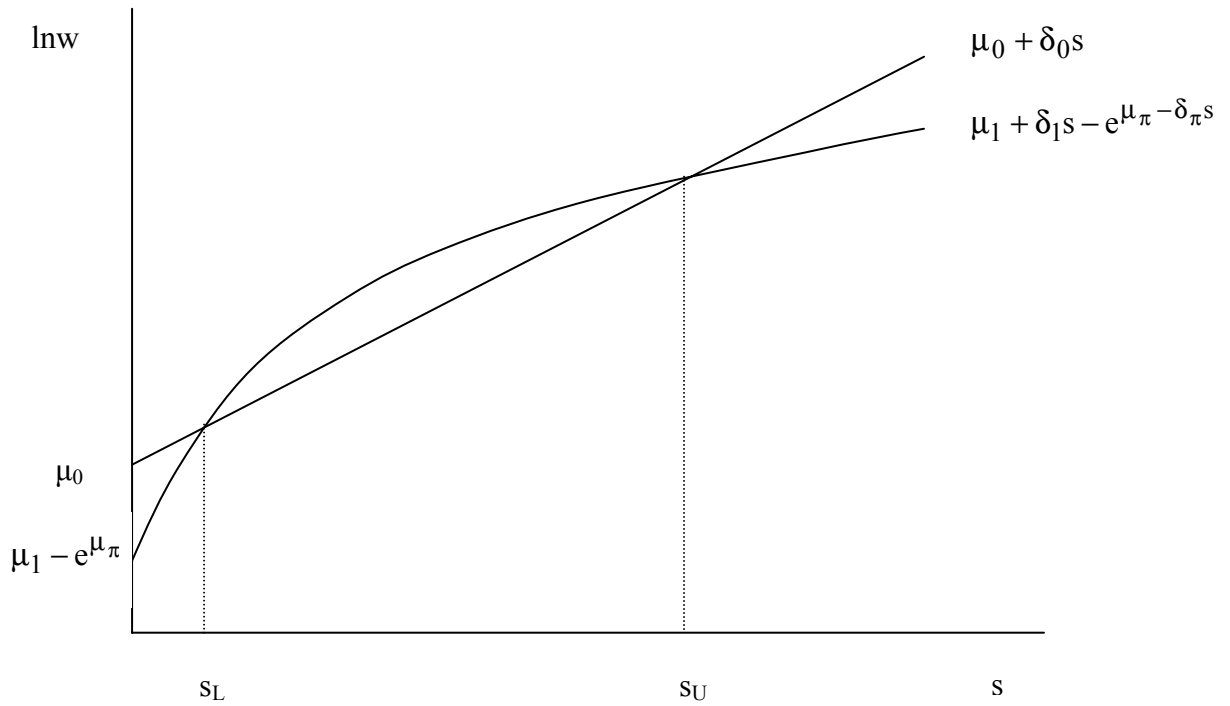


Figure 3a: U.S. and Mexico Wage Distributions 1990, Males

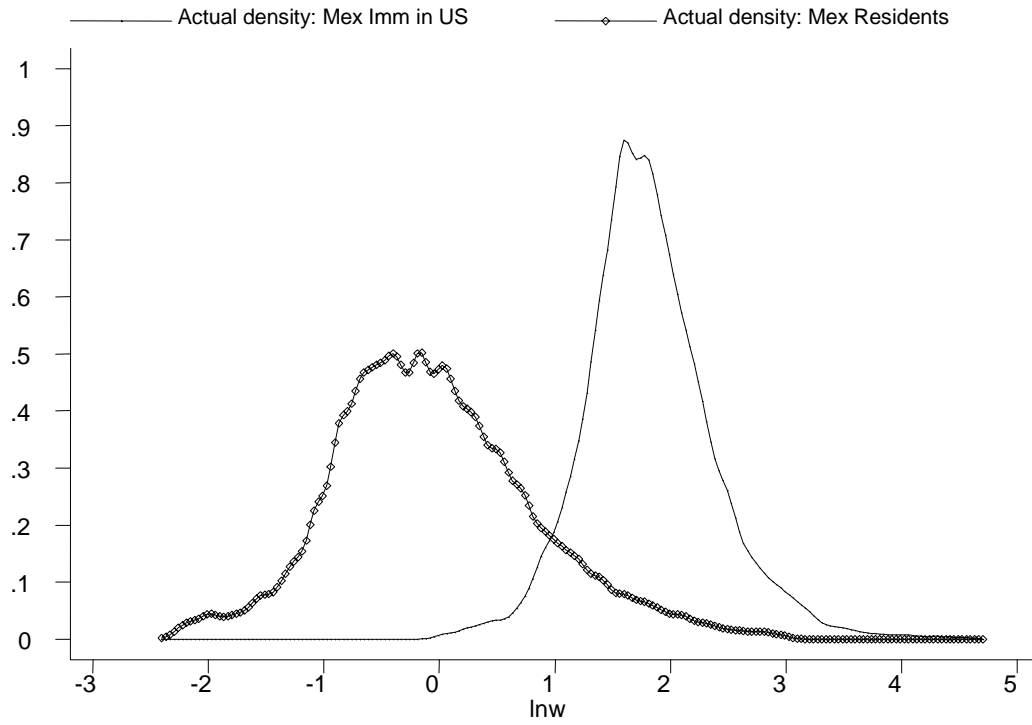


Figure 3b: U.S. and Mexico Wage Distributions 1990, Females

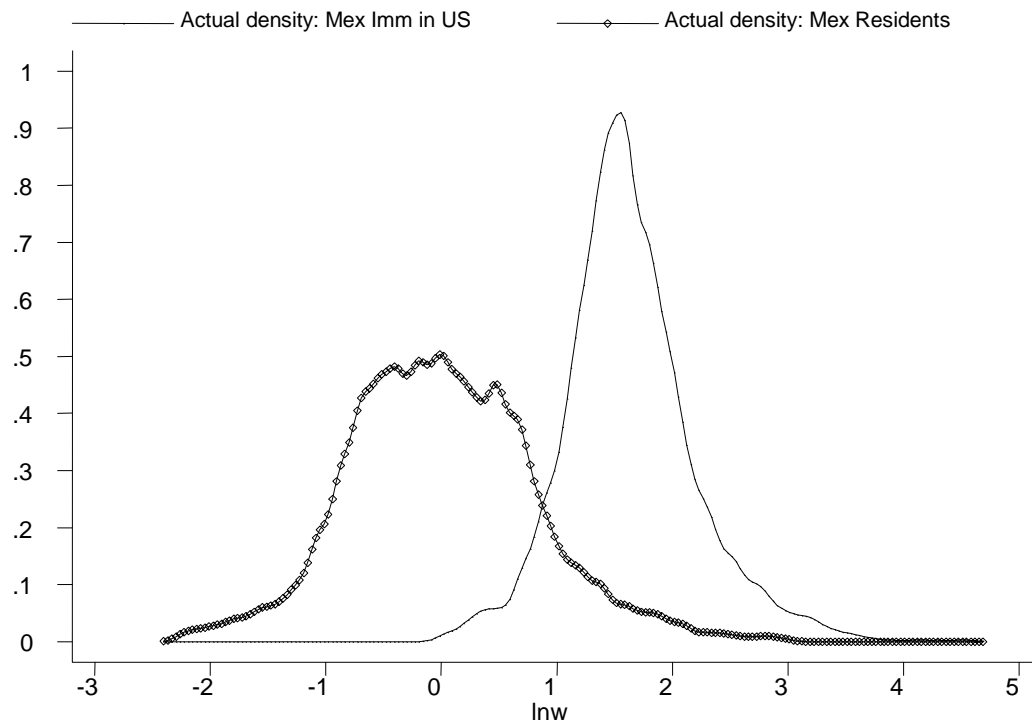


Figure 4: Predicted U.S.-Mexico Hourly Wage Differentials, 1990
(1990 U.S. dollars)

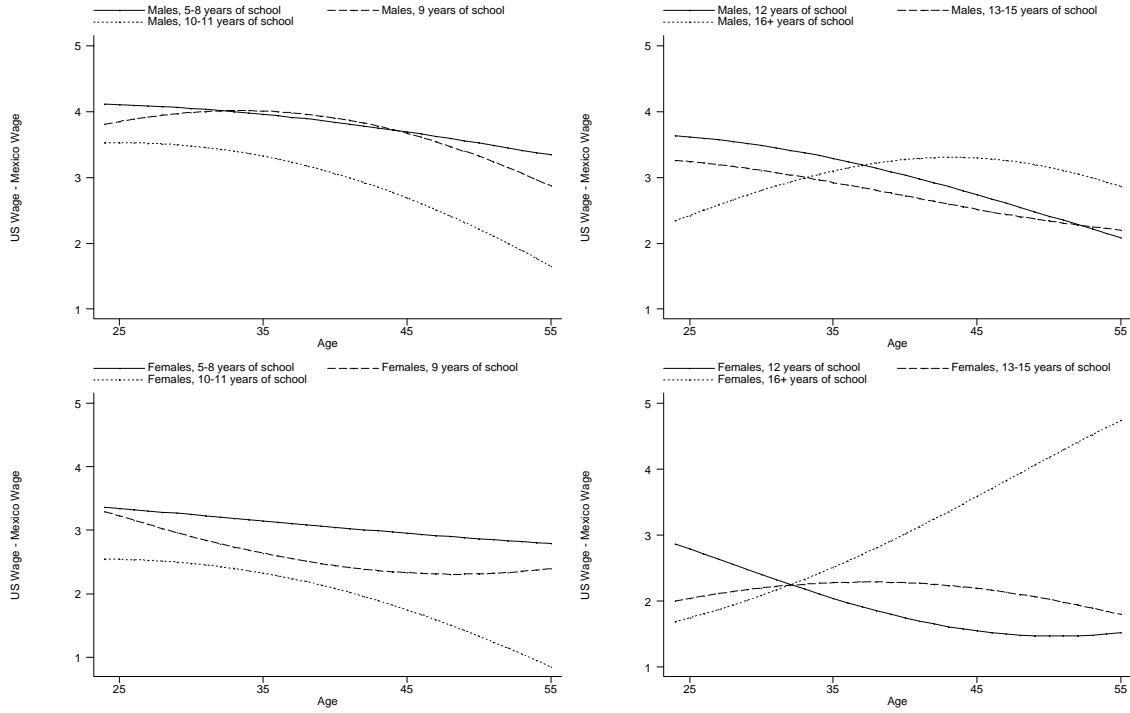


Figure 5: Predicted Probability of Migration, 1990

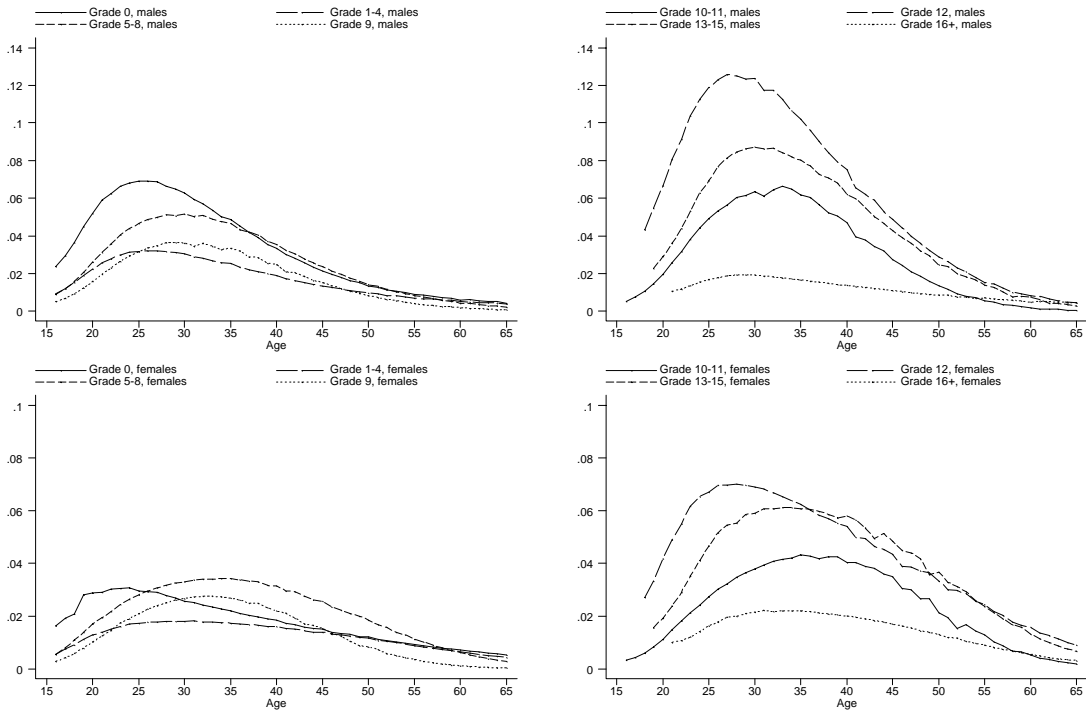


Figure 6a: Counterfactual Wage Densities for Mexican Immigrant Males, 1990
(Based on skill prices in Mexico)

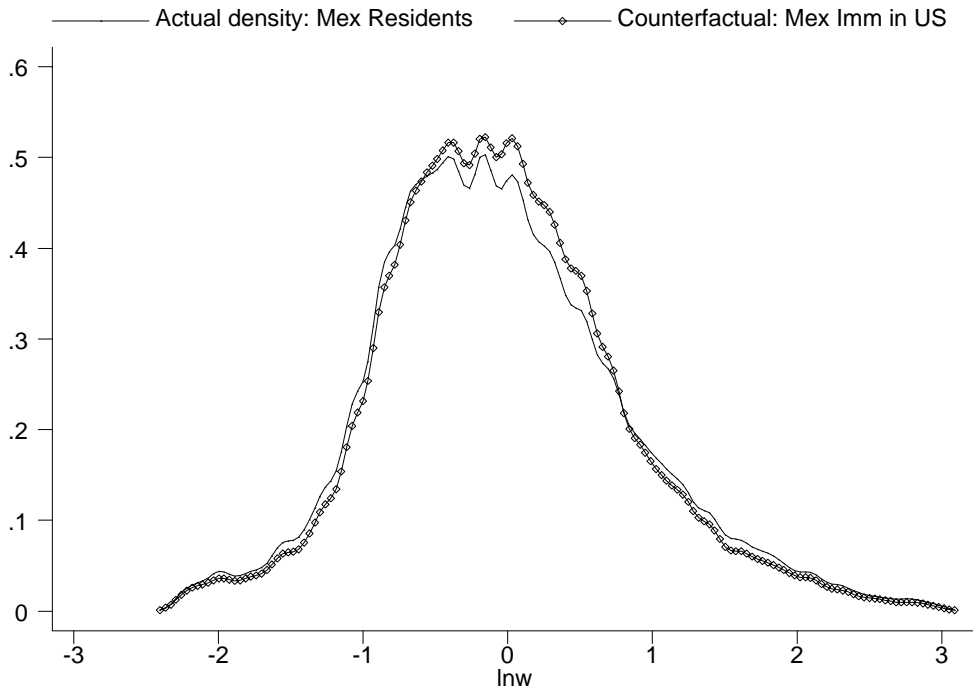


Figure 6b: Counterfactual Wage Densities for Mexican Immigrant Females, 1990
(Based on skill prices in Mexico)

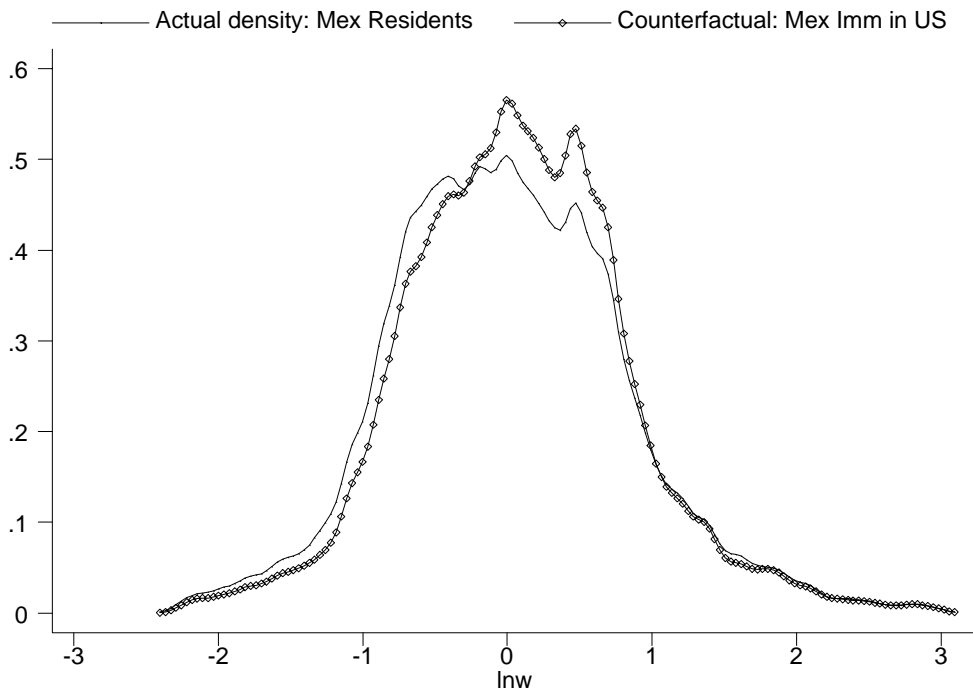


Figure 7a: Counterfactual Wage Densities for Mexican Immigrant Males, 1990
(Correcting for Undercounted Illegal Immigrants)

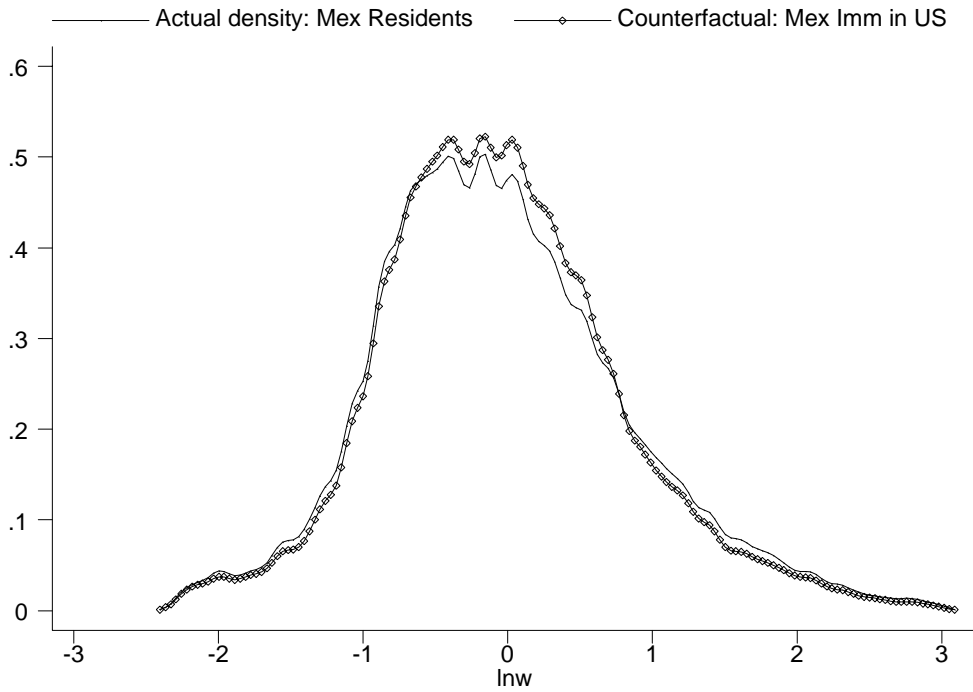


Figure 7b: Counterfactual Wage Densities for Mexican Immigrant Females, 1990
(Correcting for Undercounted Illegal Immigrants)

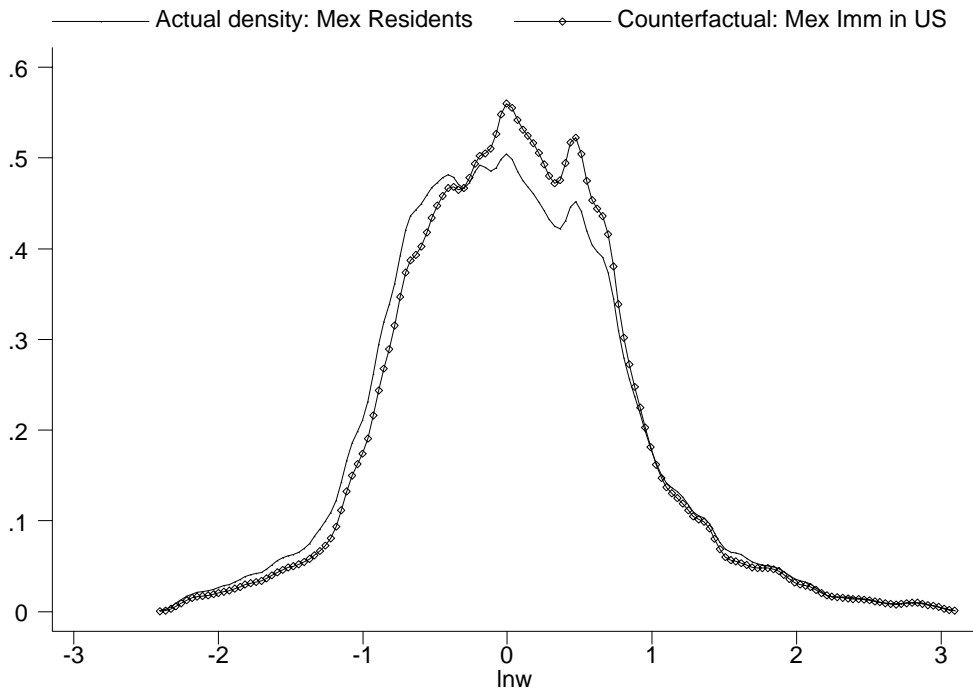


Figure 8a: Counterfactual Wage Densities for Mexican Resident Males, 1990
(Based on U.S. skill prices)

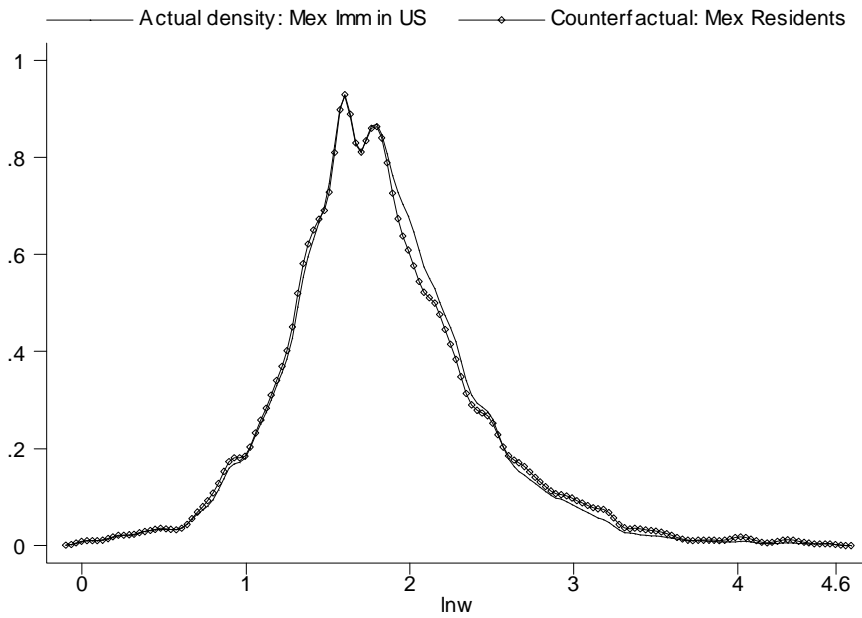


Figure 8b: Counterfactual Wage Densities for Mexican Resident Females, 1990
(Based on U.S. skill prices)

